

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCTTCAGCCCTTGTAAATTGGACATCTGCTGCTTCATATT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGTTACCTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTTATGCATTG
CTACCATTATGTTGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAACACAAGGCTGGCCTGTACTTGGAAACTGAGTTGTTAGGACTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCCTTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCCTCAT
TATATATGTTGTTCAGACCATCCTTCCTACCAAATGCAGGCCAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTGACAGTGGCAATTGGACTGATTAGAACAGAAACTCCATTGGAACCCGAGG
ACAAAGGTTATGTGCTTCACATGACTACTGCAGCAGAATGGTCTATGTCATTTCCTTCTT
GGTTTTCTGACTTACATTGATTTCTGAGAAAATTCTTACGGGTGGAAGCCAATTACA
TGGATTAACCCCTATGACACTGCACCTTGCCTATTAAACAATGAACGAACACGGCTACTTCCA
GAGATATTTGATGAAAGGATAAAAATTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGAAATTTCACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMINIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFIRLLLVIWCGVSALSMLTCSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAEWSMSFSFFGFFLTYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGCGGACGCGTGGGGAGAGCCGAGTCCCGCTGCAGCACCTGGAGAAGGCAGACC
GTGTGAGGGGGCTGTGGCCCCAGCGTGTGGCTCGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTG
GATTGGGTGGCTTCTCATGCCAATTGTTAAAGACTATGAGATACGTCAGTATGTTGACAG
GTGATCTCTCCGTGACGTTGCATTTCTGACCATGTTGAGCTCATCATCTTGAATCTTAGG
AGTATTGAATAGCAGCTCCCGTATTTCACTGAAAATGAACCTGTGTGAATTCTGCTGATCCTGG
TTTCATGGTCCTTTTACATGGTATTTATTGTGAGCAATATCGACTACTGCATAAACACGA
CTGCTTTCTGTCTTATGGCTGACCTTATGTATTCTTCTGAAACTAGGAGATCCCTTCC
CATTCTCAGCCAAAACATGGATCTTACATAGAACAGCTCATCAGCGGGTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTTCTGGATTGGTGTCAACTGCCATACACTACATGCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCACTGCTGCAAACCATGGATATGAT
CATAAAGCAAAAGAAAAGGATGGCAATGGCACGGAGAACAAATGTCAGAAGGGGAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAAGTGTACCACTTCAGCATCAGGAAGTGAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAGAATTAAAGCAGGCAGCTTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTCAAGGGAAATATTTAATTTCTTGGTT
ACTTTCTCTATTACTGTGTTGGAAATTTCATGGTACCATCAATATTGTTGATCGAGTT
GGGAAAACGGATCCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGAATCCAATTGATGT
GAAGTTTGGTCCAACACATTCTCATCTGGAAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTTACCAAGTCTTTATGCCATCTAGCAGTAAGTCTCCAATGTCATTGCTCTG
CTATTAGCACAGATAATGGCATGTACTTGTCTCTGTGCTGATCGAATGAGTATGCCCTT
AGAATACCGCACCATAACTGAAGTCCTGGAGAACACTGCAGTTCAACTCTATCACCGTGGTTG
ATGTGATCTCTGGTCAGCGCTCTCTAGCATACTCTTCTCTATTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAATTAA
GATATAAGAGGGGGAAAATGAAACCAGGGCTGACATTATAACAAACAAAATGCTATGGTAGC
ATTTTCACCTCATAGCATACTCCTCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCGTGGATATGAGGCTGG
TGTAGAGGGAGAGGGAGGCCAAGAAACTAAAGGTGAAAATACACTGGAACCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAACACGTAGGATTCGTTAAGGTTCACATGGAAAGGTTAGCTTGC
CCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAAAAAGGGCGGCCCG
ACTCTAGAGTCGACCTGCAGAAGCTGGCCGCATGGCCCAACTTGTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFFWKLGD
FPILSPKHGILSIEQLISRVGVIGVTIMALLSGFGAVNCPTYMSYFLRNVTDTDILALERRLQ
TMDMIISKKKRMMARRTMFQKGEVHNKPGFWGMIKSVTTSASGSENLTLIQQEVDALEELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQMIMGY
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGATGTCTGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAAGTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCAGCAGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
TCCATCTGGACCACGAGGCTCCTGGTCAAGGCTTTGCGTGCAGAAGAGCTTCCATCCAGGT
GTCATGAGAATTATGGGATCACCCCTGTGAGCAAAAGGCGAACAGCAGCTGAATTCACAG
AAGCTAAGGAGGCCTGAGGCTGGACTAAGTTGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGCTTGAACACTTGAGCTATGGCTGGGTTGGAGATGATTGTCATCTCTAG
GATTAGCCAAACCCAAGTGTGGAAAATGGGGTGGGTCTGATTGAAAGGTTCCAGTGA
GCCGACAGTTGAGCCTATTGTTACAACACTCATCTGATACTTGGACTAACCTGTCATTCCAGAA
ATTATCACCAACCAAAAGATCCCATTCAACACTCAAACACTGCAACACAAACAGAATTATG
CAGTGACAGTACCTACTCGGTGCATCCCCACTCTACAATACCTGCCCCACTACTCTC
CTGCTCCAGCTTCACTCTATTCACCGGAGAAAAAAATTGATTGTCACAGAAGTTTATG
GAAACTAGCACCAGTCTACAGAAACTGAACCATTTGTGAAAATAAAGCAGCATTCAAGAATGA
AGCTGCTGGGTTGGAGGTGTCACCCACGGCTCTGCTAGTGTCTCTCTCTTTGGTCTG
CAGCTGGCTTGGATTGCTATGTCAAAGGTATGTGAAGGGCTTCCCTTTACAAACAAGAAT
CAGCAGAAGGAAATGATCAGAACCAAAGTAGTAAAGGAGGAGAACGGCAATGATAGCAACCCCTAA
TGAGGAATCAAAGAAAATGTATAAAACCCAGAACAGAGTCCAAGAGTCCAAGCAAAACCTACCGTGC
GATGCCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGAACACACCTGAGGCTGGTTCTT
CATGCTCTTACCCGCCAGCTGGGAAATCAAAGGGCAAAGAACCAAAGAACAGTCCA
CCCTGGTCTTAACCTGAAATCAGCTCAGGACTGCCATGGACTATGGAGTGCACCAAAGAGAAT
GCCCTCTCTTATTGTAACCTGCTGGATCCTATCCTCCTACCTCAAAGCTCCCACGGCT
TTCTAGCCTGGCTATGTCTTAATAATATCCACTGGGAGAAAGGAGTTTGCAAAGTGCAAGGAC
CTAAAACATCTCATCAGTATCCAGTGGTAAAAGGCCCTCTGGCTGTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTACTGATTCCGAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAACACGTATCCCACCTGACATGCTCTCTGAGCCGGTAAGAGCAAAAGAAT
GGCAGAAAAGTTAGCCCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTGTAAA
GCTAAATAAAGAAATAGAACAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAAC
ACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGAATCACTGTTAGAACACACACA
CTTACTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAAATCTATAAAATTCTATTCTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAAGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAATTGCACATAGTAG
AACGCTATCTGGAGCTATTTCTGAGTTCTAGCTATCTACTTCCAAACTAAT
TTTATTTGCTGAGACTAATCTTATTCTATGCAACCATTATAACCTTAATT
TATTATTAACATACTAACAGAAGTACATTGTTACCTCTATATACCAAGCACATTAAAAGTGC
ATTAACAAATGTACTAGCCCTCTTTCCAACAAGAAGGACTGAGAGATGCAGAAATATT
TGTGACAAAAAATTAAAGCATTAGAAAACCTT

FIGURE 6

MARCFSLVLLTSIWTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNNPKCGKNGGVLIWKVPSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFFPTNKNQQKEMIETKVVKEEKANDSNPNEESKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGACCCGCGGGCCGCCACCGCGCCGCTCCGCATCTGCACCCGCAGCCCCGC
GGCCTCCCGCGGGAGCGAGCAGATCCAGTCGGCCCGCAGCGCAACTCGGTCCAGTCGGGGCG
CGGCTCGGGCGCAGAGCGGAGATGCAGCGCTTGGGGCCACCCCTGCTGTGCCCTGCTGCTGGCG
CGGCGTCCCCACGGCCCCCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAGTCAGCCGGCCG
GCTCTCACTACCCGAGGAGGCCACCTCAATGAGATGTTCCGAGGGTTGAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGGGTGAAGAGATGGAGGCAGAAGAAGCTGCTGTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTACCTCCCAGCTATCAAACTGAGACCAACACAGAC
ACGAAGGTTGGAATAATACCATCCATGTGCACCGAGAAATTCAAGATAACCAACAAACAGAC
TGGACAAATGGTCTTTCAAGAGACAGTTACATCTGTGGGAGACAGAAGAAGGCCAGAAGGG
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTGCCAGCTCCAGTAC
ACCTGCCAGCCATGCCGGGCCAGGGATGCTGCAACCGGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTCTGGGTACTGCACCAAAATGCCACCCAGGGCAGCAATGGGACATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGCTGTGCTGTGCCCTCAGAGAGGCCAGTGTCCCTGTGTGC
ACACCCCTGCCGTGGAGGGCGAGTTGCATGACCCGCCAGCCGGCTCTGCCACCTCATCAC
CTGGGAGCTAGAGCTGTGGAGCCTTGACCGATGCCCTGTGCCAGTGGCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGCAAGCCGACCTCGTGGGAGGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGTGGGGAGGCCCTGCCAGC
CCGCCGCTGCACTGCTGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCA
TAGAAAATAGCTAATTATTGCCAGGTGTGCTTAGGCGTGGGCTGACAGGCTTCTCT
CATCTCTCCCAGTAAGTTCCCCTGGCTGACAGCATGAGGTGGTGTGCATTGTCAGCT
CCCCCAGGCTGTTCTCCAGGCTCACAGTCTGGCTGGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTGCCACCCCTGTCAGATTGGCTGCTTGCCCTACAGTTGGCAGACAGCG
TTTGTCTACATGGCTTGATAATTGTTGAGGGAGGAGATGGAACAAATGTGGAGTCTCCCT
TGATTGGTTTGGGAAATGTGGAGAAGAGTGCCTGCTTGCAACATCAACCTGGAAAAATG
CAACAAATGAATTTCACGCAGTTCCATGGGCTAGGTAAAGCTGTGCCCTCAGCTGTGC
AGATGAAATGTTCTGTCACCCCTGCAATTACATGTGTTATTGACAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCAATATCCAAAGATCAATTCCCTCTCTCAGCACAGCCTGGG
AGGGGGCATTGTTCTCTGCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTTGTGACTCTAAGCTCAGTGCCT
CTCCACTACCCCACACCAGCCTGGTGCACCAAAAGTGTCTCCCCAAAAGGAAGGAGAATGGGAT
TTTTCTGAGGCATGCACATCTGAATTAGGTCAAACATTCACATCCCTCTAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGCAGCCGCTCTAATGAAGACAATGAT
ATTGACACTGTCCTCTTGGCAGTTGCATAGTAACTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAACAGTACTTAGTAATTGTAGGGCAGGATTATAATGAAATTG
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGGTGAATATGCACTGCCAACACTGAACCTACGCCACTCCAAATGATG
TTTCAGGTGTATGGACTGTTGCCACCATGATTGACAGTCTTAAAGTTAAGTTGCA
CATGATTGATAAGCATGTTCTTGAGTTAAATTATGTATAAACATAAGTGCATTAGAA
ATCAAGCATAAATCACTCAACTGCAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLNLPPSYHNETNTDKVGNNTIHVHREIHKITNNQTGQMVFSE
TVITSGVDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMLCTRSECCGDQLCVWGHC
TKMATRGNSNGTICDNQRDCQPGLCCAFQRGLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACCGTGGGGCTGTGAGAAAGTCCAATAAACATCATGCAACCCAC
GGCCCACCTTGTGAACCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTCAATCTGCAAATCTATGGGTCTGGGGCTTCTGGACCCCTT
AACTGGGTACTGCCCTGGCCAATGCGTCTCGCTGGAGCCTTGCCTCCTACTGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCTTAATCTGCCTTCATCCGACACTCCGTTACC
ACACTGGGTCAATTGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCCCCGGTCATCTG
GAGTATATTGACCAACAGCTCAGAGGAGTGAGAACCCCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGCGCCTCTGGTGTCTGGAAAAATTATCAAGTTCCTAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTCTGTCTCAGCCAAAATGCGTTCATGCTACTCATGCGAAC
ATTGTCAGGGTGGTCGTCCGGACAAAGTCACAGACCTGCTGCTGTTCTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGCTCTGCTCTTTCTCCGGTCGATCCGGGCTGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGCCTAT
GTCATGCCAGCGGCTCTCAGCGTTTCCGGATGTGTGAGCACGCTCTCCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAAACGGCTCCCTGGACCGGCCCTACTACATGTCAGAGCCTTCAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCG
CCCTGATCCAGGACTGCACCCCACCCCAACCGTCCAGCCATCCAACCTCACTTCGCTTACAGGT
CTCCATTGTGGTAAAAAAAGTTTAGGCCAGGCGCCGTGGCTCACGCCGTAACTCCAACACT
TTGAGAGGCTGAGGCCGGGATCACCTGAGTCAGGAGTTCGAGACAGCCGGCAACATGGTG
AAACCTCCGTCTCTATTAAAAACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCACTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAACAAACAAA
AAGATTATTAAAGATATTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVLDKVTDLLLFFGKLLVVGGVGVLSSSSSGRIPLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

CCCCCGCGCCCGGCCGGCGCCGGCGCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCCTGGGAGCCCTGC
TCCCTGCTCAGCTGCGCCTGCCTCTGCGCTCTGCCCTGCATCCTGTGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTCACGTTCTTCCTCTTCCCTGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGAGGAGGGGCC
GGGATCCCCACCGTCCGCAGGCCACATCGACTGTGGCTCCCTGCTGGTACCGCCTGTCTACCG
CATGTGCTCGCCACGGCGCCCTCTCTCTCTTACCCCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCAGAATGGGTTGGTTCTTAAGTCCCTGATCTGGTGGGCTC
ACCGTGGGTGCCTTACATCCCTGACGGCTCCTCACCAACATCTGGTCTACTTCGGCGTGTGG
CTCCTCCCTCTCATCCTCATCCAGCTGGTGTGCTCATCGACTTGCACACTCCTGGAACCAGCGGT
GGCTGGGCAAGGCCGAGGAGTGGCATTCCCGCTGGTACCGCAGGCCCTCTCTTACTCTC
TTCTACTTGCTGTCGATCGCGCCGTGGCGTGTGTTACTACACTGAGGCCAGCGGTGCCA
CGAGGGCAAGGTCTTCATCAGCTCAACCTCACCTCTGTGTCTGCGTGTCCATCGTGTGTCC
CCAAGGTCCAGGACGCCAGCCAACCTGGTCTGCGTGCAGGCCCTCGGTACCCCTACACCATG
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCAGGGCTATGAGACCCAGTGGTGGATGCCCGAGCATTG
TGGGCCTCATCATCTCCTCTGTGCACCCCTTCATCAGTCTGCGCTCCTCAGACCCAGGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCACCCACTATGCTAGACGCCACACAGCAGCAGCAGCAG
GGCAGCCTGTGAGGCCGGGCTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCC
ACTTCTGCCCTGGTGTGGCCTCACTGCACGTATGACGCTCACCAACTGGTACAAGCCGGTGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGGTGAAGATCTGTGCGCAGCTGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCGCGCAACCGCGACTCAGCTGAGCAGCAGCCTCA
CAGCCTGCCATCTGGCCTCCTGCCACCTGGCCTCTCGGCTCGGTGACAGCCAACCTGCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGGC
CTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCTGCGAGAGCCCCATCCCCCGCAC
ACCCACACGGTGGAGCTGCCCTCCCTCCCTCCCTGGTGTGGTGAAGATCTGTGCGCAGCTGGCAGGGCTGCT
AGGGCTCCCTGTCCCTCAGGCTCCACGGGAGCAGGGCTGCTGGAGAGAGCAGGGAACTCCCAC
TGGGGCATCCGGCACTGAAGCCCTGGTGTCCCTGGTACGTCCCCCAGGGACCCCTGCCCTC
GACTTCGTGCCCTACTGAGTCTCTAAGACTTTCTAATAAACAGCCAGTGCAGTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFLFILIQLVLLIDFAHAWNQRWLKAE
ECDSRAYAGLFFFFTLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNQNSLLQASVITLYTMFVTWSALSSIEPKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCFLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWVKICASWAGLLLWTLVAPLLLNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGCAGGCCAGGAACCACCCGTTAAGGTGTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCATGCTCTCTGCAGCAATATCCATTCCATCAACCCACACAA
CTCATGCCAGGATTGAGTCCTATGAAGGAAGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTGTTGTCACCTTGACCTCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATAATTCTCTGGCAGTTTCGATTAAAGTGTAAACTTGCATATGCTGTGCAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAAG
TGATCCTTCGAAGCTTCTCAAGGGCTTGGCTATGTGCTGCCATCTTCATTCA
CTTGCCTGGATTGAGACGTGGTCTGGATTCAAAGTGTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTCAGGATGCTTCAGAGAGGGCAGCACTTACCTGGTGGCTTCTGATG
GTCAGTTTATTCCCTCCTGAATCGAACAGGATCTGAAGAAGCTGAAGAAAACAGGACAGT
GAGAAACCACTTAGAACTATGACTACTTTGTTAAATGTGAAAACCCCTCACAGAAAGTC
ATCGAGGCAAAAGAGGCAGGCAGTGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGCTTATTGAACAGCTAACAAAGATTATTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTGCCTGCCTGTCGGCTGGTAAGGTAATGTCATGATTCACTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTTAATC
AAAAGACTTAATATAATTGAAGTAACACTTTTAGTAAGCAAGATACTTTTATTCAATTCA
AGAATGGAATTTTTTGTTTCATGTCTCAGATTATTGTATTCTTTAACACTCTACATT
TCCCTGTTTTAACTCATGCACATGTGCTTTGTACAGTTAAAAGTGTAAATAAAATCTG
ACATGTCATGTGGCTAGTTTATTCTGTTGCATTATGTGATGGCCTGAAAGTGTGGA
CTTGCAAAAGGGAGAAAGGAATTGCGAACATGTAAAATGTCACCAGACATTGTATTATT
TTATCATGAAATCATGTTTCTCTGATTGTTCTGAAATGTTCAAATACTCTTATTGAAATGC
ACAAAATGACTTAAACCATTCAATCATGTTCTTGCCTCAGCCAATTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESTYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGVLPPIISFILAWIETWFLDFKVLQPQEAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTCGGGACCCAGGACCCCTCGGGCCGACCGCCAGGAAAGACTGAGG
CCGCGGCTGCCCCGCCCCGCTCCCTGCGCCGCCGCTCCCGGACAGAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCCGCTGCTCTGCTACTGGCCCTGGGCCTGGGTGCAGGGCTGCCCAT
CCGGCTGCCAGTGCAGCCAGCACAGACAGTCTCTGCACTGCCGCCAGGGGACCACGGTGCC
CGAGACGTGCCACCCGACACGGTGGGCTGTACGTCTTGAGAACGGCATACCATGCTGACGC
AGGCAGCTTGCGGCCCTGCCGGGCTGCAGCTCCTGGACCTGTACAGAACAGATGCCAGCC
TGGCCAGCGGGGCTTCCAGCCACTGCCAACCTCAGAACCTGGACCTGACGCCAACAGGTG
CATGAATACCCAATGAGACCTTCCGTGGCTGCGGCCCTCGAGCGCCTACCTGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTCGACACGCTGACCGCCTCTGGAGCTCAAGCTGC
AGGACAAAGAGCTGCCGGACTGCCCGCTGCCGCTGCCGCCCTGCTGCTGCTGGACCTCAGC
CACAACAGCCTCTGGCCCTGGAGGCCGATCCTGGACACTGCCAACCTGGAGGGCTGCCGCT
GGCTGGTCTGGGCTGCACTGGAGGGCTCTCAGCGCTTGCGCAACCTCCACGACC
TGGATGTGCTGCCACACCAGCTGGAGCGAGTGCACCTGTGATCCGAGGCCCTGGGCTGACG
CGCCCTGCCGCTGCCGCCAACACCCGCAATTGCCAGCTGCCGCCAGGACCTGCCGCCCTGGC
TGCCCTGCAAGGAGCTGGATGTGAGAACCTAAGCCTGCAAGGCCCTGCCCTGCCGACCTCTGGGCC
TCTTCCCCGCCCTGCCGCTGCTGGCAGACGCCAACCTCAACTGCCGCTGCCCTGAGC
TGGTTGGCCCTGGGTGCGCGAGAGCCACAGTCAACTGGCCAGGCCCTGAGGAGACGCCGCTGCCA
CTTCCCCGCCAACAGCTGGCCGCTGCTCTGGAGCTTGACTACGCCGACTTGGCTGCCAG
CCACCACCAACAGCCACAGTGCCAACACAGGCCAGGAGGCCCTGGTGCAGGCCACAGCCTGTCT
TCTAGCTTGGCTCTACCTGGCTTAGCCCCACAGGCCAGGCCACTGAGGCCAGGCCCT
CACTGCCAACCGACTGTAGGGCTGTCCCCAGGCCAGGACTGCCACCGTCCACCTGCCCTCA
ATGGGGCACATGCCACCTGGGACACGGCACCACTGGCTGCTTGTCCTGGCCAGGCTTCA
GGCCTGACTGTGAGAGCCAGATGGGGCAGGGGACACGCCAGGCCCTACACCAGTCAGGCCAG
GCCACCACGGTCCCTGACCCCTGGCATGAGCCGGTGAAGGCCACCTCCCTGCCGCTGGGCTGC
AGCGCTACCTCCAGGGAGCTCCGTGCAGCTCAGGAGCCTCCGCTCACCTATCGAACCTATCG
GGCCCTGATAAGCGCTGGTGAAGCTGCCGACTGCCCTGCCCTGCCGCTGAGTACACGGTACCCA
GCTGCCGCCAACGCCACTTACTCCGTCTGTGTCATGCCCTGGGGCCGGCGGGTGCCGGAGG
GCGAGGAGGCCCTGGGGGAGGCCATACACCCCAAGCCGCTCACTCCAACACGCCAGTCACC
CAGGCCCGAGGGCAACCTGCCCTCCATTGCCGCCCTGGCGGGTGCCTCTGCCGCG
GCTGGCTGCCGCTGGGGCAGCTACTGTGTCGGCGGGGCCATGGCAGCAGCGGCTCAGG
ACAAAGGGCAGGTGGGCCAGGGCTGGGCCCTGGAACCTGGAGGGAGTGAAGGTCCTTGGAG
CCAGGCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
ACTCATGGCTTCCCAGGGCTGGCTCCAGTCACCCCTCCACGCAAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGGGCTCTCAGCCAGTGAGATGCCAGGCCAGGCCCTCTGCTGCC
ACACCACGTAAGTCTCAGTCCAAACCTGGGATGTGTCAGACAGGGCTGTGACCAAGCT
GGGCCCTGTTCCCTCTGGACCTCGTCTCTCATCTGTGAGATGCTGCTGCCAGCTGACGCC
CTAACGCCCCAGAACCGAGTGCTATGAGGACAGTGCTCCGCCCTGCCCTCCGCAACGTGCA
CCTGGGACGGCGGGCCCTGCCATGTGCTGTAACGCACTGCCCTGGTCTGCTGGCTCTCCAC
TCCAGGGGGACCCCTGGGGCCAGTGAAGGAAGCTCCGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGACTCTAGTCTGGGCCAGGAAGCGAAGGAACAAAAGAAAATGGAAAGGAAGATGC
TTTAGGAACATGTTTCAAACCTCAGAGACAAGGACTTGGTTTGTAAAGACAAACGATGATATGAA
GGGAAGATGTTTCAAACCTCAGAGACAAGGACTTGGTTTGTAAAGACAAACGATGATATGAA
GGCCTTTGTAAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLPLLLLLALGPGVQGCPGQCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLHSNSLLALEPGILDANVE
ALRLAGLGLQQLDEGLFSRLRNLIHDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLALPGDLSGLFPRLRLAAARNPFCNCVCPPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVFPQDCPPSTCLNGGTCHLGRHHLACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRSLTGLIEPVSPSLRVGLQRYLQGSSVQLRSILRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPGLPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAYCVRRGRAMAAAQDKGQVPGPAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFPGPGLQSPLHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGGCAGGGCGGGTGGCTGAGTCCGTGGCAGAGGCGAAGGCACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGTGCGGTGCTGAGCTTGGCCTGGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTGCACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTTGAATTCAAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCTCCAAATCCAGAAAACAGGACTATGAAGAGCCAAGAAGTA
CGGAAACCGAGCTTGACCGCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTCT
TTTCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAGACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGAATGAAAATCCTTAATGGAAG
CAATAAGAAAAGCaaaaAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATGCTCTTTATTGGTGAATTCTGCCACAGAAATATCCAG
GCAGCGAGAGAGATGTTGAGAAGCTGACTGAGGAAGGCTCTCCAAAGGGACAGACTGCTCTGG
CTTCTGTATGCCCTGGACTGGTGTAAATTCAAGTCAGGCCAAAGGCTCTGTATATTACAT
TTGGAGCTTGGGGCAATCTAATAGCCCACATGGTTGGTAAGTAGACTTTAGGGAGGCT
AATAATATTAACATCAGAAGAATTGTGGTTATAGCGGCCACAACCTTTCAGCTTCATGATC
CAGATTGCTTGTATTAAGACCAAATATTCAAGTGAACCTCCTCAAATTCTGTTAATGGATAT
AACACATGGAATCTACATGTAATGAAAGTGGTGGAGTCCACAATTTCCTTAAATGATTAG
TTTGGCTGATTGCCCTAAAAAGAGAGATCTGATAAAATGGCTCTTTAAATTCTCTGAGTTG
GAATTGTCAGAATCATTTTACATTAGATTATCATAATTAAAATTTCCTTAGTTTCA
AAATTTGTAATGGGGTATAGAAAAACACATGAAATATTACAAATTTGCAACATGC
CCTAAGAATTGTTAAATTCACTGGAGTTATTGTGCAGAATGACTCCAGAGAGCTACTTC
TTTTTACTTTCATGATTGGCTGTCTCCATTATTCTGGTCAATTGCTAGTGACACTGT
GCCTGCTCCAGTAGTCTCATTTCCATTGGCTAATTGTTACTTTCTTGCTAATTG
AAGATTAACATTTAATAAAATTATGCTAAGATaaaaaaaaaaaaaaaaaaaaaaa
aaa

FIGURE 18

MRVRIGLTLILLCAVLLSLASASSDEEGSQDES LDSKTTLSDESVKDHTAGR VVAGQI FLDSEEESEL
ESSIQEEEDSLKSQEGESVTEDI SFLESPNPNENKDYE EEPKVKRPALTAIEGTAHGE PCHFPFLFLDK
EYDECTS DGDREDGRLWCATTYDYKADEK WGFCETEEEAAKRRQM QEAEMMYQTGMKILNGSNKKSQKR
EAYRYLOKAASMNH KALERV SYALLFGDYL PONIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYT FGA LGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAAGATTTAAGCCCATTCTGCAGTGGAAATTCACTGAACAGTAGCAAGAGGACACCATTCTCTT
GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTTTGGGTGCTAGG
CCTCCTAATCCTCTGGTTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGCTTGGAAACTTGGCAGCCAGAACCTTAAAGGCAGA
AAAAAGGGATTTCATGTAACTCGCTGTGACTGAATCAGGATCAACAGCTTAAAGGAGA
AACCTCAGAGAGACTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAACGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTGGGAGAAAGGTCTCTGGGTGCTGATCAATAATGCTGGTGT
CCCAGCGTGCTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATGCTTCCTTGGCAAGAAAGCTCAAGGGAGAG
TTATTAAATGTCTCCAGTGTGGAGGTCGCCTGCAATCGTGGAGGGGCTTATACTCCATCCAAA
TATGCAGTGGAAGGTTCAATGACAGCTTAAGACGGGACATGAAAGCTTGGGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAAACTGGCAGATCCAGTAAAGGTATTGAAAAAAAC
TCGCCATTGGAGCAGCTGTCTCCAGACATCAAACAAAATGGAGGTTACATTGAAAAAA
AGTCTAGACAAACTGAAAGGCAATAAACTTGTGAACATGGACCTCTCCGGGTAGAGTG
CATGGACCACGCTTAACAAAGTCTCTCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTCTGGATAACCTGTTCACATGCCAGCAGCTTGAAGACTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAACCCAAGGCAGTGTGACTCAGCTAACCAACAAATGCTCCCCAGGCTATGA
AATTGGCGATTCAAGAACACATCTCCTTCAACCCATTCCTATCTGCTCCAACTGGACT
CATTTAGATCGTGTTTGGATTGCAAAAAGGAGTCCCCACCATCGCTGGTGGTATCCAGGG
CCCTGCTCAAGTTTCTTGAAAGGAGGGCTGGATGGTACATCACATAGGCAAGTCCCTGCC
GTATTTAGGCTTGCCTGCTGGTGTGTGAAGGAATTGAAAGGACTTGCCCATTCAAAATGA
TCTTTACCGTGGCCTGCCCATGCTTATGGTCCCAGCATTACAGTAACTGTAATGTTAAGT
ATCATCTCTTATCTAAATTAAAAGATAAGTCAACCCAAAAAAAA
AAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGGLISVTLNMLPLVKKAQGRVINSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIKEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGGCGCGGTAGCATGGAGGGGGAGACTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGGGC
CACTCGCTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTCTTCTGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGGTGTTATACAATTGA
CATTCAAGAAATATATTCCATGCTATCAGCTTTAGCTTTATAATTCTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATTCAAGATCAGATCATGACGTTAGAGAGAGGCTGCTTCACAAAAACTGCAGGAGCATT
TTCAAAACCAAGACCTGTTTCTGCTATTAACACCAAGTATAAACAGAAAGCTGCTCTACTC
ATCGACTGGAACATTCTTATATAACCTCAAAAGGACTTTACAGGGTACCTTAGTGGTT
GCCAATCTGGGCATGCTGAACAACGGTTATAAAACTGTATCAGGTTCTGTATGTCACACTGG
TTTAGCCGAGCAGTACAAACACAGCTCAAATTTTGAAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAAACTAGTAAAGGATGTAACAGATTAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAAGGAGCAAGAGAGAAGAACATCAGGAGAACCTCAGGAGA
ACATTTTCTTGTCAAGGCATTACGGACCTTTTCCAAATTCTGAATTCTCATTCACTGTGTT
ATGTCTTAAAAAATAGACATGTTCTAAAGTAGCTGTAACACTACAACCACATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAAGCATAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATCTGGTAGTAGTAACCAAGATAAACGATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTGGTGAATTACGGCTC
CTACATTTGATCTTTAACCTACAAGGAGATTTTTATTGGCTGATGGGAAAGCCAAAC
ATTTCTATTGTTTACTATGTTGAGCTACTGCACTAGTTCAAGTCTTACTATGTTACC
TGTTTGAGCTAATACACAGATAACTCTTAGTCATTCAACAAAGTACTTTCAACATCA
GATGCTTTATTCCAAACCTTTTCACTTCAACTAGTTGAGGGAAAGGCTACACAG
ACACATTCTTGAATTGGAAAAGTGAGACCAAGGACAGGACAGTGGCTCACACCTGTAATCCACAG
TAGGGAGACAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCAGTCTATTAAAAAATGGAAAAGCAAGAATAGCCTTATTTCAAAATATGGAAA
GAAATTATATGAAAATTATCTGAGTCATTAACCTCCTTAAGTGATACTTTTAAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDEGFLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVGSCMSTGFSRAV
QTHSSKFFEDGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDDRWFQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGAGGGCAGAGTCAGCCGAGCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCAA
GCAGCGCGCAGCGAACGCCGCCGCCACACCCCTGCGGTCCCGCGCGCTGCCACCCCTCCCTCCCC
GCGTCCCCGCCCTCGCCGAGTCAGCTTGCGGGTTGCGTGCCTCGCCGAAACCCGAGGTCAACAGCCCGCCTCT
GCTTCCCTGGGCCGCGCCGCCCTCACGCCCTCCCTCTCCCTGGCCCGCCTGGCACCGGGGACCGTTGCCTGA
CGCGAGGCCAGCTACTTTGCCCCGCGTCTCCGCCCTGCGCTCTCCACCAACTCAACTCTTCTCC
TCCAGCTCCACTCGCTAGTCCCCGACTCGCCAGCCCTCGGCCGCTGCGTAGCGCCGCTTCCGTCGGTCCAAA
GGTGGGAACCGTCCGCCCGCCCGCACCATGGCACGGTTCGGCTTGCCCGCCTCTGCACCCCTGGCAGTGC
AGCGCCGCCGCTGCGCTGCCGAGCTCAAGTCGAAAAGTGCTCGGAAGTGCAGCTTTACGTCAAAGGCTTC
AACAAAGAACGATGCCCTCCACGAGATCAACGGTGTCAATTGAAGATCTGTCCCAGGGTTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGAAAGTAAAGATGATTTCAAAGTGTGGTCAGCGAACAGTGCATTTG
CAAGCTGTTGCTTCACTTACAAGAAGTTGATGAAATTCTCAAAGAACTACTGAAAATGAGAATCCCTG
AATGATATGTTGAAAGACATATGGCATTATACATGCAAATTCTGAGCTATTAAAGATCTTCGCTAGAGTTG
AAACGTTACTACGTGGTGGAAATGTGAACCTGGAAGAAATGCTAAATGACTCTGGCTCGCCCTGGAGCGATG
TTCCGCTGGTGAACTCCAGTACCACTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACTGGAGCAGCTGAAG
CCCTCGGAGATGCCCCTGCAAATTGAAGCTCCAGGTTACTCGTGTGTTGTAGCAGCCCGTACTTCGCTCAAGGC
TTAGCGGTTGCGGGAGATGTCGTGAGCAAGGTCTCGTGTGACTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTTTGGCCAACCAAGGGATCTGATTTGAATGAAACATTCTAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCTTCAACATTGAATCGGTATGGATCCCATCGATGTGAAGATTCTGATGCTATTGAAACATGCAGGAT
AATAGTGTCAAGTGTCTCAGAAGGTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTCTCGT
TCCATCTGAAAGTGCCTCAGTGTGCTCGTCAACCATCACCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTGGACCACTGGTTACTGATGTCAGGAGAAACTGAAACAGGCCAAGAAATTCTGGCTCCCTCCGAGCAAC
GTTTGCAACGATGAGAGGATGGCTGAGGAAACGGCAATGAGGATGACTGTTGAATGGAAAGGCAAAAGCAGGTAC
CTGTTGCACTGACAGGAAATGGATTAGCCAACCAGGGCAACAAACCCAGAGGTCCAGGTTGACACCAGCAAACGAC
ATACTGATCTCGTCAAATCATGGCTTCTCGAGTGTGACGAGCAAGATGAAAGATGCATACAATGGGAACGACGTG
GACTCTTGATATCAGTGTGAAAGTAGTGGAGAAGGAAGTGGAGTGGCTGTGAGTATCAGCAGTGCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGAGAGTGCACATGAGAAAGGCCACAGTGTGTTGTGCTGGTCTGGGCA
CAGGCCACTCCTCAGTGTGATCTGCACTTGTCTGGTATGAGAGAGTGGAGATAATTCTCAAACCTCTGAG
AAAAAGTGTTCATAAAAAGTTAAAGGCCACCAAGTTATGACTTTTACCATCCTAGTGTACTTGCTTTAAATGAA
TGGACAACAATGTACAGTTTACTATGTCGCTGGCACTGGTTAAGAAGTGTGACTTTGTTCTCATTGAGTTGGG
AGGAAAAGGGACTGTGCAATTGAGTTGGTCTGCTCCCAAACCATGTTAACAGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTCATTGATTTTACTCTATTATGTTGTTGATGTTTTCTCATTGTTGTGGTT
TTTTTTCCAACGTGTGATCTGCCATTGTTCTTACAAGCAAACCGGGTCCCTTGGCACGTAACATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTATTATCATGTTATCTTATTAAAGAAAAGCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKSCEVRRLYVSKGFnKNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSILNDMFVKTG
LYMQNSELFKDLFVELKRYYVGPNLEEMILNDFWARLLERMFRVLVNSQYHFTDEYLECVSKY
TEQLKPFQGDVPRKLKLQVTRAFVAARTFAQGLAVAGDVSKSVNNPTAQCTHALLKMIYCSHC
RGLVTVKPCYNYCSNIMRGCLANQGQLDFEWNNFIDAMILMVAERLEGPFNIESVMDP
IDVKISDAIMNMQDNSVQVSQKVFQGCQPKPLPAGRISRSISESAFSARFRPHPEERPT
TAAGTSLDRLVTDVK EKLKQAKKFWSLPSNCNDERMAAGNGNE
DCWNGKGKSYLF
AVTGNGLANQGNNPEVQV
DTSKPDILILRQIMALRVM
TSKMKNAYNGNDV
FFF
DIS
ESS
GEGSG
GCEY
QQCP
SEFD
YNAT
DHAG
KSANEKADSAGVR
PGAQAY
LLTVFC
CILFL
V
M
QREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glycans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCCATTATATTCTCAAGCAACT
TACAGCTGCACCGACAGTTGCGGATGAAAGTTCTAATCTCTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTCTAGGAGATGGCTCCAGGAAGGCCAAGAATGTGAGTGCAAAGATTGGTTCCCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCTGTGATCATT
TCAAGGCAATGTGAAGAAAACAAGACACCAAAGGCACACAGAAAGCCAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGCTCTGCCTTTGTAGGAGCTCTG
AGGCCCACTCTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAAGACACTC
TTCTTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAGCA
TGTTTTCAAGATCATTTGTTGCTCTCTAGTGTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCCTAACCCAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT
AGCTAGTGTCAATTAAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTT
AAATGTCAAAAAAAAAAAAAAA

FIGURE 26

MKVLISSLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPDCDFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCTGCAGAGGCTGAGCAGGGAAAAGCCAGTCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCACCTCCTGCAGCTGCTGGTCTGCTTCTTACCTGCC
CCTGCACCTCATGGCTCTGCTGGCTGCTGGCAGCCCTGTGAAAAGCTACTTCCCTACCTGA
TGGCCGTGCTGACTCCAAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTTCAAGCAG
ATAAAGGGCTTACAGGAGCCTCGGGAAAGTGGCCACTGGAGCTGGCTGCGGAACCGGAGC
CAACTTCAGTTCTACCCACCGGGCTGCAGGGTACCTGCCTAGACCCAAATCCCCACTTGAGA
AGTTCCGTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTGTGGCTCC
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTG
CTCTGTGCAGAGCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTG
TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCTTCATGTGGCAGCAAGTTTC
GAGCCCACCTGAAACACATTGGGATGGCTGCTGCCTCACAGAGAGACCTGGAAGGATCTGA
GAACGCCAGTTCTCGAAATCCAATGGAACGACAGCCCCCTCCCTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGAAAGGCTGTCAAACAATCTTCCAAGCTCCAAGGCACCTATTGCTCC
TTCCCCAGCCTCAAATTAGAACAGCCACCCACCAGCCTATCTATCTTCACTGAGAGGGACTA
GCAGAACATGAGAGAACATTCATGTACCACCTACTAGTCCCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTCAATCCGCCTCGACAGTGAAAAGCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGATCCTCAACTGCAAGTTCTGGACTAGTCTCCAACGTTGCCTC
CCAATGTTGTCCTTCTCGTCCATGGTAAAGCTCTCGCTTCTGAGGCTACAC
CCATGCGCTCTAGGAACGGTACAAAAGTCATGGTGCCTGCATCCCTGCCAACGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTCGAGCTGGGGACCAGGGAGAATCAGAGATGCTGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLTLPLHLMALLGCWQPLCKSYFPYILMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDDVVVCTLVLCSVQSPRKVLQEVRRVLRPGGVLFFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSSKALICSFPSL
QLEQATHQPIYLPRLRT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATTTGCCTATCCACCTCCCCAAGCCCCTTAACCTATGCTGCTGCTAACGCTGCTGCTGCTGCTGCTGCTGCTAAAGGCTCATGCTGGAGTGGGACTGGTCGGGCCAGAAAGTCTCTCTG
CCACTGACGCCCCCATCAGGGATTGGGCCTCTTCCCCCTTCCTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGGAGAAAGTGGGGATGGCTAGAAAGCTGGGAGATAGGAAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGCTGCCTTATTAAA
GTGGTTGTTATGATTCTTACTAATTACAAAGATATTAAGGCCCTGTTCATTAAGAAATT
GTTCCCTCCCCGTGTTCAATGTTGTAAGATTGTTCTGTGTAATATGTCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSSAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTCGAATTCTTCAACTATACCCACAGTCACAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCCTTATTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTGATAAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGAAT**CATGTCGG**
GAAGAGATAACAATCCTGGCCTGTGTATCCTCGCATTAGCCTGTCTTGGCCATGATGTTTAC
TTCAGATTCATCACCACCCCTCTGGTTCACATTTCATTGGTTATTGGGATTGTTGTT
TGTCTGCGGTGTTTATGGTGGCTGTATTGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAATATGAAGTGCCTGCTGGGTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTGATTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTTCTCTGGGTCCCTGGGTGGCTGTGCTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGAATATAAGCCCCTTCGGGCATTGGTACATGTGGTGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTCGCTGCCAGCAAATGACTATAGCTGGGCAG
TGGTTACTGTTATTCAACAGAAAGTAAAATGATCCTCCTGATCATCCCATCCTTCGTCTCTC
TCCATTCTCTTCTACCACATCAAGGAACCGTTGTGAAAGGGTCAATTAAATCTCTGTGGTGTAG
GATTCCGAGAATCATTGTCATGTACATGCAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTCGATGCTGACTGTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACCGAGAATGCATATACTACAACGTCTATTAAATGGGACAGATTCTGTACATCAGCAAAGATGC
ATTCAAATCTGTCCAAGAACACTCAAGTCACTTACATCTATTAACTGCTTGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGGGTGTGTTCACTGTTGGAGGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGTGGCAGTCCCTCTGTTATTGGTAGCTTTTGCTACTTAGTAGC
CCATAGTTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCTGTGTTGATGTC
TGGAAACAAATGATGGATCGTCAGAAAGCCCTACTTATGGATCAAGAATTCTGAGTTGTA
AAAAGGAGCAACAAATTAAACAAATGCAAGGGCACAGCAGGACAAGCAGTCATTAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGA**TAGATACCCATTAGGTATCTGTACCTGGAAAACATT**
TCCTTCTAAGAGCCATTACAGAAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATT
TTAAAAGACCTAATAAACCCATTCTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLLVHIFISLVLGLLFVCGVLWWLYDYTNDSLIE
LDTERENMKCVLGFAIVSTGITA VLLVLIFVLRKRIKLTVELFQITNKAIS SAPFLLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMMQNALKEQQHG
ALSRYLFRCYC CFWCLDKYLLHLNQAYTTAINGTDFCTS AKA DFKILSKNSHFTSINC FGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNA RAAQD KHS LRNEEGTELQAI VR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCTCTGAGAAGAAGAGAAAAGGTTCTGGACCTCCCTGTTCTTCCTTAGA
ATAATTGTATGGGATTGTGATGCAGGAAAGCTAAGGAAAAGAATATTCAATTCTGTGGT
GAAAATTTTGAAAAAAAATGCCCTCTCAAACAAGGGTGTATTCTGATATTATGAGGAC
TGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTCCTGTTGCTGGTACTGGAGTAC
ATTCAAACAAGAAACGCAAAGAAGATTAAAGGCCAAGTTCACTGTGCCTCAGATCAACTGC
GATGTCAAAGCCGAAAGATCATGATCCTGAGTTCATGAAATGTCAGCAGGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTGATGCATCCTACTCCAGTGTGTCGGCCTGCCG
TACACAGTGGTGTGCTTGATAATTCAAGGAGGAAAATACTTGTGCGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATGTTATCCCTACACGATGGAGAGAATCCTT
TATCGTCTAGAAAGTAAACCCAAAAGGGTGTAACTACCCATCAGCTCTACATACATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACGTGACAGCCGGTCACTCTGATGCAGCTCTGGCTGTCACTGTAGCTGTGCCACCCAC
CACCTGCCAAGGCCATCCCTCTGCTGCTTCTACCACCGATCCCCAGACCAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCACTGCCACCTACACAAGCAGCCAAAAGGGCC
AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCCTCCAGAAACCTGTTGGAGC
GGATGTCAGCCTGGGACTTGTCCAAAAGAAGAATTGAGCACACAGCTTTGGAGGCCAGTATCCC
TGGGAGATCCAAACTGCAAATTGACTTGTGCTTTAATTGATGGGAGCACCAGCATTGGCAA
CGCGATTCCGAATCCAGAAGCAGCTCCGGCTGATGTCGGCCAAAGCTTGCACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCAGTATGGAGACAACCCCTGCTACTCACTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGGAGGAGCTTCTAAT
GTAGGTCGGGCCATCTCTTGACCAAGAACATTCTTCCAAAGCAATGGAAACAGAACGG
GGCTCCAATGTGGTGGTGTGATGGTGGATGGCTGGCCACGGACAAAGTGGAGGAGGCTTCAA
GACTTGCAGAGAGTCAGGAATCAACATTCTTCATCACCATTGAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGGCCAACTTGCAAAACAAGGCCGTGTCAGAACAAACGGCTTACTC
GCTCCACGTGCAAGCAGCTGGTTGGCCTCCACAAGACCCCTGCAAGCTTGTGAACTCGGCTGACATTGGCTCGTACGAC
GGCTCCAGCAGTGTGGGGACGGCAACTTCCGCACCGTCTCCAGTTGTGACCAACCTCACCAA
AGAGTTGAGATTCCGACACGGACACGCGCATCGGGGGCGTGCAGTACACCTACGAACAGCGG
TGGAGTTGGGTTCCGACAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGC
TACTGGAGTGGTGGCACCAGCACGGGGCTGCCATCAACCTCGCCCTGGAGCAGCTTCAAGAA
GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCAGCAGGGAGGTCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCCATCTGAAGGGAGTGTACACCTATGCGATAGGCCTGGCTGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTCCCCGCCAGAGACCACCTTCTTGTGGACGA
GTTTGACAAACCTCATCAGTATGTCCTCAGGATCATCCAGAACATTGATCAGAGTTCAACTCAC
AGCCTCGGAACTGAATTCAAGAGCAGGCAGAGCACCAGCAAGTGTGCTTACTAAGTGTGTT
GGACCACCCACCGCTTAATGGGCACGCACTGGTGCATCAAGTCTGGCAGGGCATGGAGAAC
AAATGTCATTGTTATTCTTGCCATCATGCTTTTCAATTCCAAAATGGAGTTACAAGA
TGATCACAAACGTATAGAATGAGCCAAAGGCTACATCATGTTGAGGGTGTGGAGATTTCAT
TTTGACAATTGTTCTAAAATAAATGTCGAATACAGTGCAGCCCTACGACAGGTTACGTAG
AGCTTTGAGATTAAAGTGTATTCTGATTTGAACTCTGTAACCTCAGCAAGTTTCAT
TTTGTCATGACAATGTAGGAATTGCTGAATTAAATGTTAGAAGGATGAAAAAATAAAAAAA
AAG
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLSEPKKGVTYPSALTYSSKSPAAQAGETTKAYQRPIPPTTAQPVTLQMLLAVTVAVA
TPPTTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNCCCCMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTTLQPLVKRVCDTDLACSKTCLNSADIGFVIDGSSSSVGTGNFRTVLQFVTN
LTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGWYWSGGTSTGAAINFALEQ
FKKSKPNKRKLMILTDGRSYDDVRIKPAMAHLKGVITYAIGVAWAQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCCTTAGGAGGTGGCTGCCTGTTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTTCTGTTTCAGAGTAGTTCAACACAGATCTGAGTGTAAATTAAGCATGGAA
ACAGAAAACAACAAAAACTTAAGCTTAATTTCATCTGGAATTCCACAGTTCTAGCTCCCTGGACCC
GGTGACCTGTTGGCTTCCCGCTGGCTCTATCACGTGGTCTCCGACTACTCACCCCGAGTGTAA
AAGAACCTTCGGCTCGGTGCTCTGAGCTGCTGGATGGCCTCGGCTCTGGACTGTCCCTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCTGCTGCTGTACTCCTGAGTTCTTGATGTGGTAC
CTCAGCCTCCCCACTACAATGTGATAGAACCGGTGAACGGATGTACTCTATGAGTATGAGCGATTAA
CAGACAAGACTTCACTTCACACTCGAGAGCATTCAAACATGCTCTCATCAAATCCATTCTGGTCATT
TGGTGACCTCCCACCCCTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTAAAAAGTCT
TGGTGGGGATATGAGGTTCTTACATTTCATTAGGCCAAGAGGCTGAAAAGAAGACAAAATGTTGGC
ATTGTCCTAGAGGATGAACACCTCTTATGGGACATAATCGACAAGATTTTAGACACATATAATA
ACCTGACCTTGAAAACCATTATGGCATTCAAGGTGGTAACTGAGTTTGCCTTCAAGTACGTAATG
AAGACAGACACTGATGTTTCAAAACTGGCAATTAGTGAAGTATCTTAAACCTAAACCACTCAGA
GAAGTTTCACAGGTATCCTCTAAATTGATAATTTCCTATAGAGGATTACCAAAAAACCCATATT
CTTACCAAGGAGTATCCTTCAAGGTGGTCCCTCCACTGCAGTGGGTGGTTATATAATGTCCAGAGAT
TTGGTGCCAAGGATCTATGAAATGATGGTCACGTAACCCATCAAGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATTAAAAGTGAACATTCAATTCCAGAACAGACAAATCTTTCTTCTATATAGAATCC
ATTGGATGTCGTCAACTGAGACGTGTGATTGCAGCCATGGCTTTCTCCAAGGAGATCATCACTTT
TGGCAGGTCACTGCTAAGGAACACCACATGCCATTATTAACTCACATTCTACAAAAGCCTAGAAGGACAG
GATACTGAAACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTATTAGTCAGG
CCCTTCAAAGATGATGTGGAGGAATTAAATAAAGGAATTGGAGGTTTGTCAAAGAAATTAAATAGG
ACCAAACAATTGGACATGTCATTCTGTAGACTAGAATTCTAAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGAAAAACAAACATGTAGAGTTATTGAACAAATGTAGTCATTGAAGGTTGTAA
TATCTTATGTGGATTACCAATTAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATA
CTGAACAAAATTTCACCTGTTGGTCATTATAAGTACTTCAGATGTCAGTGTGCTGGCT
ATTATTAAAATTACTCAACTTGTGTTTAAATGTTGACGATTCAACAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTTCAGTTACTTAACAGTTGATCAGTTATTGATACATCACTCCA
TTAATGTAAGTCATAGGTCAATTGCAATATCAGTAATCTTGGACTTGTAAATATTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNCNSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIHQDFLDTNNNLKTTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

FIGURE 38

MELGCWTQLGLTFLQLLISSLPREYTVINEACPGAEWNIMCRECCYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSPCFHDTGTCVLDKAGSYKC
ACLAGYTQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQQNGEWSGKQPICIKACREP KISDLVRRVLPMQVSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPGMYQHLHTQLQYECISPFYRLGSSRTCLRTGKWSGRAPSCIPI CGKNIENITAP
KTQGLRWPWQAAIYRRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDREDEKTIQSLQISAIILHPNYDPILLADIAILKLLDKARI STRVQPI CLAASR
DLSTSFQESHITVAGWNVLADVRSPGFKN DTLRGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPS DICTAETGGIAAVSFPGRASPEPRWHLMGLVWSYDKTC SHRLSTAF TKVLPFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTCCTACATCCTCTCATCTGAGAACAGAGAGCATATCTTACGGGCCGTGATTATTAAACGTGGCTTAATCTGAAGGTTCTCAGTCAAATTCTTGATCTACTGATTGTGGGGCATGGCAAGGTTGCTTAAGGAGCTTGGCTGGTTGGCCCTTGTAGCTGACAGAAGGTGGCAGGGAGAATGCAGCACACTGCTCGAGAATGAAGGCGCTCTGTTGC
TGGTCTTGCCTTGGCTCAGTCCTGCTAATCATTGACAATGTGGCAACCTGCACCTCCTGTATTCAAACCTCTGTA
AAGGTGCCCTCCACTACGGCTGACCAAAGATAGGAAGAGCGCTCACAAAGATGGCTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCCAGAGGTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCTAGACA
ACCTGCCTACGTGTCCTCGCAGAGGACGGCAGGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAAGTA
GGCACGCCCTTGAGAGATCCACTATTAGACGAGATCATTAAAAAAATCGAGCTTGGAGTGTCTTCGA
GGACAAAGAGCGGGAGTGCAGTTGCAACCATGCCGACCAGGGCAGGGAAATTCTGAAAACACCAGTGCCCTGAAG
TCTTCCAAGGTTGTAACACCTGATTCCAGATGGTAAAATTACACGATCAAGATCAATCGAGTAGATCCCAGTGA
GCCTCTCTATTAGGCTGGTGGAGGTAGCAGAACCCACTGGTCCATATCATTATCCAACACATTATCGTGA
TGATGCCAGAGACGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCACCGGATGGACATCAGCAATGTCCTC
ACAACCTACGCTGTGCGTCTCTGCCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTAACAGAAGTCCGCA
GCAGGAACAATGGACAGGCCCGGATGCCCTACAGACCCAGATGACAGCTTCATGTGATTCAACAAAAGTAGCC
CCGAGGAGCAGCTGGATAAAACTGGTGCAGGCTGGATGAGCTGGGTTTCATCTCAATGTGCTGGATGGC
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGGCCATCAATGGACATGATCTCGATATGGCA
GCCAGAAAGTGCAGCTCATCTGATTAGGCCAGTGGAAAGACGTGTTACCTCGTGTGTCACGCCAGGTCGGCAGC
GGAGCCCTGACATCTTCAGGAAGCCGCTGGAACAGCAATGCCAGCTGGTCCCAGGGCAGGGAGAGGAGCAACA
CTCCCAAGCCCTCCATCCTACAATTACTGTCATGAGAAGGTGTTAAATCTCAAAAGACCCGGTGAATCTCTCG
GCATGACCGTGCAGGGGAGCATCACATAGAGAATGGGATTGCTATCTATGTCATCAGTGTGAGCCGGAGGAG
TCATAAGCAGAGATGGAGAATAAAACAGGTGACATTGTTGAATGTGGATGGGTCGAACGTGACAGAGGTCA
GGAGTGAGGAGCTGGCATTATTGAAAAGAACATCCTCGATAGTACTCAAAGCTTGGAAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACACATGCCCAACCCAGTGA
GGTCATGTGGCTGGAAATTACACGGTCTGTATAACTGTAAGATATTGTTACAGAAGAAACACAGCTGGAAAGTC
TGGGCTTGCATTGTTAGGAGTTATGAGAACATCAAATCCATTGTTGAAGGAA
CACACAGCATACAATGATGAGAACATTAGATGTGGTGAATTCTCTGCTGTCATGGTAGAAGTACATCAGGAATGA
TACATGCTGCTGGCAAGACTGCTGAAAGAACCTTAAAGGAAGAACATTACTCTAATGTTCTGGCTGGCACTT
TTTATAGAATCAATGATGGTCAGAGAAAAGAGAAAATCACAAATAGCTAACAGTGAACACTATATTATC
TTGTCAGTTTTATATTAAAGAACATCATTGTAAGGAAAAGTATGTCAGGAAAGTATGATCATCTAATGAAAGCCAGT
ACACCTCAGAAAATGATTCCAAAAAAATTAAAGAACACTAGTTTTCTCAGTGTGGAGGATTCTCATTACTCTAC
AACATTGTTATATTCTTATTCAATTTAAAGCCCTAAACAACTAAATGATTGTTGATACCCACTGAATT
CAAGCTGATTAAATTTAAATTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGCCATTTTAATTACAGCT
AAAATTTTAAATGCTGAGAACGTTGCTTCATCAAACAGAACATAATTTCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTATAPS
PEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGSRNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNMGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYPRDDSFHVLINKSSPEEQLGIKLVKDEPGVIFNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRSPDFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDRGR
IKTGDILLNVDGVELTEVRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLRLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAAGGCATTGTATCTTCAGTTGTCATCAAGTTGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACAGTAATAAAAACATGGGCTTCAACCTGACT
TTCCACCTTCTACAAATTCCGATTACTGTTGCTGTGACTTGTGCTGACAGTGGTTGGGTGGC
CACCAAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTAAAGCAAAGGAGTTCATGGCTAATTCC
ATAAGACCCCTCATTTGGGAAGGGAAAAACTCTGACTAATGAAGCATTCCACGAAGAAGGTAGAACTT
GACAACGTGCTCTGTGTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTCAAACCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGTAAAG
CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGAACAGAGAGAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCTACCCAGGCTGAAGG
TAAAAAGTTAATCGAGCAAACCTTGAATGTGGCTATCTAGAACGCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAAATGACTTTAACCTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTG
GGGTGTTACTGCCCTAACGAGAGCAGTTCAAGGTGAATGGATTCTCTAACAACTACTGGGAT
GGGGAGGCGAACAGCATGACCTCAGACTCAGGGTTGAGCTCCAAAGAACGAAATTCCCGGCCCTG
CCTGAAGTGGTAAATATAACATGGCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTTACACCAAGTGTACGAGCTGGAGAACAGATGGGTGAGTAGTTGTTCTTATAAAT
TAGTATCTGTGGAACACAATCCTTATATATCAACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTGGTGAATGTTGGAAGAACTGATTCTTGTGCAATAATTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTGAGCTGAATTTCCTTTGTATTTCT
TAGCAGAGCTCCTGGTGAATGTAGAGTATAAACAGTTGTAACAAAGACAGCTTCTTAGTCATTTGAT
CATGAGGGTTAAATATTGAATATGGAACTTGAAGGACTTATATAAAAGGATGACTCAAAGGATAA
AAATGAACCGCTATTGAGGACTCTGGTGAAGGAGATTATTAAATTGAAAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAACTGTTATTCACTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTCAAGT
GAGAAGGCGTCCAAAAGAGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGAGGACCTGCCAGGTATGCCCTCCAGTGATGCCACCAGAGAACATCTCTATTAGT
TTTAAAGAGTTTGAAATGATTGTACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFIFHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALESREQFFKVNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGACTGAGCTCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTCCTGCCAGCCTGACCAGTGGCTCTGTTTCCCACAACAG
ACGGGACAACCTGCAGAGCTGCAACCCAGGACAGAGCTGGAGGCCAGGGCCAGCTGGATGCCAT
GTTCCAGAGGCCAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCGTCCCCTCCC
TTCCTTATTATTCTGCTGCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTGTT
TCCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLILASLTGSVFPQQTGQLAEIQLPQDRAGARASWMPMFQRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGGTGGTGGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCACACAAACCCCTCTTGTACCCATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACGTAAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACCGAGCACCTGCAAAGCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGCAAGCAGCCAATGAGTCCCATAATGGTCCATCCTCCCCATCTC
CTGGAGATGGGAGAAAGTGATATGACCTTCATCTGCCTGCCAGGAACCTGTCAGCAGAAACT
TCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCTCTGTTGGTGCCCTCCTGCTCAGTCTCTTGACTGGGCTATTCTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGATGGACATTGTCGG
AAACTCTAACATATGCCCTATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAACTCTAAAGGAAGATCCAGCAAATACGGTTACTCCACTGTGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACCTCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAASGPVKELVGSVGGAVTFPLKSKVQVDSIVWTFNTTPLVTIQP
EGGTIIVTQNRNRERVDFPDGGYSLKLSKLKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLICLLVPILLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATAACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCTCAAAACCCATCTTGCTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTCTTCATCATTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTTAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTCATTGAAAAACATCAGTGACATTCACTCCAGAATCCTCAA
CTTGCAGTGGTTTCAATGACTCTTGTGCACCTCCTACTGGTTCAATAAACCCACAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTCTCAGTATTTTAGGTCTATTGCTTGGAAATTCTGGAGGTCTGTTGGGCT
CAGTCAGATAGTCATCGGTTCCCTGGCTGTGTGGAGTCTAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGGAATAAAATGTAAGTATCAGTAGTTGAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPIMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCACTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGAGAGAGTGACCCCTGGCCCTCTCCTACTGGCAGGCCTGACTGCCTTGGAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCAGGGCTCCTGGCATTGCTGGATCGCGCAGTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACAGCACAGTCCTGTACCTGAGAAGGCCATCCCACACTCATCACTCCAGGCTCTGCCA
CTACTTGCTGAGCACAGGACTGCCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCCTGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTATGAATTAAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALLEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGAAGGAGGAGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGAATGAAGTTCCAG
GGGCCCCCTGGCCTGCCCTGCTGGCCCTGCTGGCAGTGGGAGGCTGGCAGAGA
CGGAGAGGAAAGCACTGGGACAAATATTGGGAGGCCATTGGACATGGCCTGGAGACGCCCTGA
GCGAAGGGTGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGGCAAGGGACAGAGAAGCAGTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGCGC
AGCAGATGCTTGGCAACAGGGTGGGAAAGCAGGCCATGCTCTGGGAAACACTGGGACAGAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGTGCTGCCGGCTCTGGCAGGG
GTGCCTGGCCACAGTGGCCTGGGAAACTCTGGAGGCCATGGCATCTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATACCCG
GAAACTCAGCAGGCCAGTTGAATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTGGACCAACACTCAGGGAGCTGGCCAGCCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAAATGAAGGGTCACGAATCCCCCACCCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGCTCACAGTGGCAGCAGTGGCAGTGGCAGCAATGGTACAACAAC
AATGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGTGGCAGCAGCAGTGGCAGCAG
TGGCGGCAGCAGTGGCAGCAGTGGCAGCAGTGGCAGCAGCAGTGGCAGCAG
CCTCCTGGGATCCAGCACGGCTCCTCCGGCAACCACGGTGGAGCGGGGAGGAAATGGA
CATAAACCCGGGTGTGAAAGCCAGGGAAATGAAGCCCGGGAGCAGGGAAATCTGGGATTCAAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATGCCCTTGG
GAGGCTGGAGACAATTATGGGGCAAGGGTCAGCTGGGAGTGGCAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCATACTGTGAACTCTGAGACGCTCCTGGATGTTAACTTGCACACTTCTGGAA
GAATTTAAATCCAAGCTGGTTCATCAACTGGGATGCCATAAACAGGACCAGAGAAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACAGATTGGATGGAGGCCACACTCCCTCTTAA
AACACCAACCTCTCATCAACTCAGCCCTGCCCTGAAATAACCTTAGCTGCCCAACAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 52

MKFQGPLACLLALCLGSGEAGPLQSGEESTGNTIGEALGHGLGDALEVGKAGAIGKEAGGAAGSKVSEALGQGTREAVGTGVRQPGFAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGDADAVRGSWQGPVGHSGAWETSGGHGIFGSQGGLGGQGNGPGLTPWVHGPNSAGSGFMNPQGAPWGQGGNGGPPNGTNTQGAVAQPGYGVSVRASNQNEGCTNPPSGSGGGSSNSGGSGSQSGSSGSGSNGDNNNGSSSGGSSGSSGSSGGSSGGSGNSGGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRLLGGSGDNYRGQGSSWGSGGDAVGGVNTVNSETSPGMFNFDTFWKNUFKSKLGGFINWDAINKDORSSRIE

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGGACAAGCTGCCGACAGAAGGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCTCAGACCGGTGGCAATGCCCCATGGCTACTCCTGCTGCTGGTTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTGGACCTATGCCCTCTATAACAACTGCCGCCGGCTCCAGTGTTCC
CACAGCCCCAAACGGAACCTGGTTGGGTACCTGGGCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAAGGCTTACGGTATGGCTGGGTCCCAT
CATCCCCCTCATGTTTATGCCACCCGTACACCATCCGGTCTATCACCAATGCCCTAGCTGCCA
TTGCACCAAGGATAATCTCTCATCAGGTTCTGAAGCCCTGGCTGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTGGATGCTGACGCCGCCCTCCATTCAACATCCT
GAAGTCCTATATAACGATCTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGCTGGACATGTTGAGCACATCAGCCTCATGACCTGGACAGTCTA
CAGAAATGCATCTCAGCTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTGTAGAGAAAAGGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGCGGCCTCCACAGGGCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTCATCCGGGAGCGCGCGTGCACCCCTCCCCACTCAGGGTATTGATGATTTTCAAAGACAA
AGCCAAGTCCAAGACTTGGATTTCATTGATGTCCTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATAAAGAGCAGAGGCTGACACCTTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGTCTGTACAACCTGCGAGGCACCCAGAAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGACGACCTGGCCC
AGCTGCCCTTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTTGCGGCGATCCTGAGGTCTACGACC
CCTTCCGCTTGACCCAGAGAACAGCAAGGGAGGTACCTCTGGCTTTATTCCCTTCTCCGCA
GGGCCAGGAACTGCATGGGCAGGCAGGCGTCCGACAGGAGATGAAAGTGGTCTGGCGTTGAT
GCTGCTGCACTTCCGGTTCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTTGCAGTGACTTCTGAC
CCATCCACCTGTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWLGRLPVAMSPWLLLLLVGSWLLARILAWTYAFYNNCRLQCFPQPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGLLGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SILMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFTDAVIRERRRLPTQGIDDDFKDKAKSKTLDFIDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLRHPEYQERCRQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHNPTVWPDPVEYDPFRFPENSKGSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGCTGTTGTTGCACTTACCCGTGTTCTGCCT
TTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTCGAGTCTTGGCATTGACGTGG
TACAGCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTCTTG
ATAATTCATGGCCAGTTTATGAAGCTTGGAGGCACTATGGACAGAAGCTGGTGGACAGTTT
GTAACTATCTTCGAAACCTCTGCTTACAGACATGTGCCTTTATCTTGCAGCAATGTGTTGCCTT
GTGATTCGAACATTGAGGGTTACTTTGGAAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAAATCTCCTCATGTACCTGTTCC
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGCAGTGGAGCACCCAGCAGGCCGCCAACATGCTCTGTCTGTGCCGT
TACGTGCCGTATCGGGGAAGGCCAGACCGAGTTCACTACTTTGAGTCGAAGGGGCTCCCTGCCAGCTGAAGTCC
ATTTCAAGCTCAGTGTCTCATCCCCCTCCCAAGGAATTCTCCACCTACCGCCAGTGGAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGCAGCTAGACTTGAAGAATTGTCCATTATCTCAAGATCATGAGAAGAGCTGAGG
CTGGTGTAAAGATTGGACAAAAGATGAGGACCCATTGACCCGAGGAGATCATGCACTCCGTGGGGACTTG
GGAGTCAGATATCTGAACGAGGAGGAGAAAATTCTCAAGAGCATGGATAAAAAGCCACGATGACCATCGACTGG
AACGAGTGGAGAGACTACCACTCCATCCCCGTGAAAACATCCCGAGATCATCTCACTGGAAGCATTCCACG
ATCTTGATGTGGGTGAGAATCTAACGGCTCCGGATGAGTCAGCTGGAGGAGGGCAGACGGGGATGTGGTGAGA
CACCTGGTGGCAGGGGGTGGGGCACGGGGCTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCCTCCGCAGCAACAACTGGGCATCGTTGGTGGCTCACTCAGATGATTGAGAAGGGAGGGCAGG
TCACCTGGGGGCAATGGCATCAACGCTCTAAATTGCCCGAATGCCATCAAATTATGGCCTATGAGCAG
ATCAAGGCCCTTGTGGTGTGACCAAGGAGACTCTGAGGATTCAAGAGGAGCTTGTGGCAGGGCTTGGCAGGGCC
ATCGCCCAAGCAGCAGTCAACCAATGGAGGTCTGAAGACCCGGATGGCCTGCCAGACAGGGCAGTACTCAGGA
ATGCTGGACTGCGCAGGAGATCCTGCCAGAGGAGGGGGCTTCAACAAAGGCTATGTCCCCAACATGCTG
GGCATCATCCCCATGCCGCATCAGCCTTGAGCTACAGAGACCTCAAGAATGCCCTGGCTGCAGCACTATGCACTG
AACAGCGGGACCCGGCTTTGTGCTCTGGCTGTGGCACCATGTCCAGTACCTGTGGCAGCTGGCAGCTAC
CCCCCTGGCCTAGTCAGGACCCGGATGCCAGGCAGCAGCCTCTATTGAGGGCCTCGAGGTGACCATGAGCAGCCTC
TTCAAACATATCCTGGCAGGGGGCCCTGGGCTGACAGGGGCTGCCCTGGGCAACTTCATGAGGTCATCCCA
GCTGTGAGCATCAGTCAGTGGCTCACAGAAGCTGAAGATCACCTGGGCTGCAGTCGGGTCAGGGGGGAGGGC
GGCCGGCAGTGGACTCGTGTACCTGGGCCAGCCTGGGTGTCAGCCATCTCAATTCTGTGAAATGTGCCAACACT
AACGTGTCAGGCAAGCTGTGAAAACCTAGACGCCACCCGAGGGGGGGGGAGAGCTGGCAGGCCAGGGCTT
GTCTGCTGCCCTAGCAGGCCCTCTGGTGGTCCAGCGAAGACACAGGATTCTTAGGGTCCAGGGTCAGCAGG
CTCCGGGCTCACATGTGAGGAGGACATTTCAGTGGCTCACAGTGGCTCAATAGTGAGCTGGAGCCTGGAGGGCT
TAGTTCTCCATTTCACCCCTGCAGCCAGTGGGCCAGGCCCCCTGCCCTGTGCTGCCGTGCATCTCCCTGTGC
CCTCTGCTGCCCTGCTGTGAGGTAAGGTGGAGGGCTACAGCCACATCCCACCCCTGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCTCCACGGCTGACTTCCAAACCTACAGCATTGACGCCAACCTGGC
TGTGAAGGAAGAGGAAGGGATCTGGCTTGTCAGGCATCTGAGCCCTGCTGATGGCTGGGCTCTGGGCT
CTTGGGAGTCAGGGGCTCGGGCTCCCTGGCTCACAGAGGCAAGTGTGGGCTCATGGTCTCTGAGCT
GGCCTGGACCCCTGTCAGGATGGGCCCCACCTCAGAACAAACTCAGTGGCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTGAGGGCGAGGGCAGAGGCTTGTGTTCTGGGGAGGGCTACAGCCACATCCCACCCCTGTCCAATCCC
ACTGTTGGGAAAAGGGTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTGGCTGACTGTCAGACTGTGTCAGCTGGGACCTATTATTTGATGGCTGGGCTCTGGGCT
GGGACCAAGCCCAATTCCACTTGTGTCAGCTGGGACCTATTATTTGATGGCTGGGCTCTGGGCT
AACTATTATGATGTTAAATTAGCTGTCATTTCAGGTGTCATTTCAGTTTATTCAATTATGTTCATGGTT
GATTGACCTCCAAAGGGCCCCAGGGGATGGGAGGGAGGAGGAGGGGGCTTGGGCTGCACTCACATT
GTCCAGAGAAATTCTTTGGGACTGGAGGAGGAGAAAGGGCCAGAAGGAGCAGCAGCCCTGGCTCCCTTGGAG
GTTGGGGAGGGCTGCCCTAGGATTCAGGGTTGACTGGGGCTGGAGAGAGAGGGAGGAACCTCAAT
AACCTTGAGGTGGAATCCAGTTATTTCTGCGCTGCCAGGGTTCTTCACTCTTCTGAAATGTCAAGGAG
TGAGGTGCCCTCTCACTGTGAAATTGTGTTGGGGGGGGCTGGAGGAGGGGGCTGGCTCCCTCCAGC
CTTCTGCTGCCCTTGTCTAACATGCCGCAACTGGGACCTACGGGTTGCACTTCATTCCACCAATGACCTGA
TGAGGAATCTCAATAGGATGCAAGATCAATGCAAAATTGTTATATGAACATATAACTGGAGTCAGTCAAAAG
CAAATTAAAGAAAGAATTGGACGTTAGAAGTGTCAATTAAAGCAGCTTCTAATAAAAGTTGTTCAAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVIFPSQEFSTYRQWKQKIVQAGDKDLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFVVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAGGCAGCGGAGCTCCACTCAGCCAGTACCGAGATA
GGCTGGGAACCTTCCCCAGCC**ATGGC**
TTCCCTGGGCAGATCCTCTTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGAGGATGGAATCCTGAGCTGACTTTGAACCTGACATCAAACTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTTGGTCCATGAGTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTAGAGGCCGGACAGCAGTGTGCTGATCAAGTGTAGTT
GGCAATGCCTTTGGCTGAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTGGGTGTGAGGCTCCCCGATGGTCCCC
CAGCCCACAGTGGCTGGGCATCCAAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTCTGTGCTCTACAATGTTACGA
TCAACAAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAACATGGAGATCAAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGTTCTGTGT
CTCTTCTTCTTGGCCATCAGCTGGCACTCTGCCTCTCAGCCCTACCTGATGCTAAAT**TAAT**
GTGCCTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCAC
CACCAAGATATGACCTAGTTTATATTCTGGGAGGAAATGAATTCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAGAAGCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTCAAAGACATATTAGAAGTTGGGAAAATAATTCTGTAAGTGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCAAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGAGTGAGAGGACAGGATAGTCATGTTCTGTCTGAAATTAGTTATATGTC
TGTAATGTTGCTCTGAGGAAGCCCCTGAAAGTCTATCCAAACATATCCACATCTTATATTCCAC
AAATTAAAGCTGTAGTATGTACCTAAGACGCTGCTAATTGACTGCCACTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGTCAAATGATTCACTTTTATGATGCTTCAAAGGTGCCTGGCTTC
TCTTCCAACTGACAAATGCCAAAGTTGAGAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCGATTTATAAATAAAACTGAGCACCTTCTTTAAACAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVGLVHDFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRRSHLQLLNKASL
CVSSFFAISWALLPLSPYLMK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAACCAATGGCCAGCTATCCTTACCGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCTCGGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCTGGGGTGGTTATGGGGTCTGCCCTGGAGGGCCTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGATGTTCCCTCTGGAACTCCAGGAGGACCATATG
GCGGTGCAGCTCCCCGGGGCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCGCATCGATGTCTACGGCTCTCAGCCCTGTGAAATTCCAGCAGTGGAAAGAACCTTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCATTAGCTACACAGAGCTGCAGCAAGCTCTGCCC
AAATGGCTACAACCTGAGCCCCAGTTCACCCAGCTCTGGCTCCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGCAACATCCGGCTCAGCTCGAGGACTTCGTCA
CCATGACAGCTCTGGATGCTATGACCCAAACATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTTCCGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCTGCATCATGCCACCA
AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGGCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCTGATGGCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCGTATGCCAGTGGTGAGTGTTCATCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTCTCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACAGTGGCTTGGATTCTGCCACACCCATAAAT
CCTTGGCCAGGCTCTGCCCTGCAGCTGGACCCCTACTTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTACCTCTCCCTGCCAATACTTTTTAATTGCAATTTC
ATTGGGGCAAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYPGPPNSGGQYGSGLPPGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSTGPGPYGGAAPGGPYGQPPPSSYGAQQPGLYQGGAPPNVDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSAWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGG**A**TGCAGGGCCCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTTTGCCCGAGCATCGGTGTTGGAGGAGAAAGTTCCAAAACCTCGGGACC
AACTTGCCCTCAGCTCGGACAACCTCCTCACTGGCCCTCTAAGCTGAACATCCGAGCCGC
TCTGGACCCCTAGGTCTAATGACTTGGCAAGGGTTCCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGAGGTGGCCTCCATCGTGGGGCTGCCTGCCATG
GATTCCCTGGCCCCCTGAGGATCCTGGCAGATGATGGCTGCTGCCGTAGGACCGCCTGGGG
AGCGCTGCCTGAAGAACTCTTACCTCTCAGTGCTGCCCTCGCTCCGGCAGTGGCCCTT
TGCCTGGGAGTCTCTCCCGATGCCACAGGCCTCAGCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAATTCACTGGGAGCCGGGGAAAAATCCTTCCAAACG
CCCTCCCTGGTCTCTCATCCACAGGGTCTGCCTGATCACCCCTGGGTACCTGAATCCCAGTG
TGTCCCTGGGAGGTGGAGGCCCTGGACTGGTGGGAACGAGGCCATGCCACACCCCTGAGGGA
ATCTGGGTATCAATAATCAACCCAGGTACAGCTGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGAAATATTAATCGGTATCCAGGAGGAGCTGGGAAATATTAATCGGTATCCAGGAG
GCAGCTGGGAAATATTCACTACCCAGGTATCAATAACCCATTCCCTGGAGTTCTCCGC
CCTCCTGGCTTCTGGAACATCCCAGTGCTGGCTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGG**C**AGCACGATAGAGGGAAACCCAACATTGGAGTTAGAGTCTGCTCCGCCCTGCTG
TGTGGGCTAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCCACTTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCAATTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPAGGSAVQRWPPSWGLPAMDSWPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLGESSPDATGLSPEASLLHQDSESRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVWGGGGPGTGWGTRPMHPPEGIWGINNQPPGTWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCTTGTCCCTCCTTGCACCTCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATCCTCCTATGAGTCCAGCCTCCTGGAA
TTGCTTGAAGGAGCTGCCTCCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTCAACACCATGTTGTCGCAACACATTGACAGCCATTGAAGCCTGTGTCTTCTGGCC
GGGCTTTGGGCCGGGATGCAGGAGGCAGGCCCGACCCGTCTTCAGCAGGCCACCC
CTGAGTGGCAATAAAATTCGGTATGCTG

FIGURE 66

MGSGLPILVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGGGC
CAGGTGCCCGTCGAGGTGCCCTGGCCGGAGATGCGGTAGGAGGGCGAGCGCGAGAACCCCC
TTCCTCGGCCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCGCTGGGCGAGCCTGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTGCCTTCATCCACCAGCTCCAGCTCCAGTGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTTGGCTGCCTTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGCAGCTTGGGAGAACGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGAAGAAGGTACTTCAA
GGCAGTGCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGAAGAAGGTACTTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSSSTSSSDGNLRPEAITAIIVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GGCAGGAATAACTAGAGAGGAACAATGGGGTTATTCAGAGGTTTCTCTTAGTTCTGTGCCTGCTGCACCAAG
TCAAATACTTCCCTCATTAAGCTGAATAATGGCTTGAAGATATTGTCATTGTTATAGATCCTAGTGTGCCAGAA
GATGAAAAAAATAATTGAACAAATAGAGGATATGGTGAACAGCTTACACGTTCTACCGTACCTGTTGAAGCCACAGAAAAAAGA
TTTTTTTCAAAATGTATCTATATTAACTCTGAGAATTGGAAGGAAATCCTCAGTACAAAAGGCCAAAACATGAA
AACCATAAACATGCTGATGTTATAGTTGACCCACCTACACTCCAGGTAGAGATGAACCATACACCAAGCAGTTACA
GAATGTGGAGAGAAAGGCGAATACATTCACTTACCCCTGACCTTACTTGGAAAAAAACAAAATGAATATGGACCA
CCAGGAAACTGTTGTCATGAGTGGCTACCTCCGGTGGGGAGTGTGTTGATGAGTACAATGAAGATCAGCCTTC
TACCGTCTAAGTCAGGAAACAAAGGTGTTCCGCAGGTACTCTGTTAGAAAATAGAGTTATAAGTGT
CAAGGAGGAGCAGCTCTTAGAGGAGTGCAGAATTGCTACAAACAAAATGTGAAAGATTGTCAATTCTTT
CCTGATAAAAGTACAAACAGAAAAGCATCCATAATGTTATGCAAGATATTGATTCTGTTGAATTGTAACGAA
AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAGTGCATTTAGAAGTACATGGGGAGTGTATTAGAAT
TCTGAGGATTAAAAACACCATACCCATGGTGCACACCCCTCCACCTGCTTCTCAATTGCTGAAGATCAGTCAA
AGAATTGTTGCTTAGTTCTGATAAGTCTGAGGAAACAGCTGGGGTAAGGACCGCCCTAAATCGAATGAATCAAGCAGCA
AAACATTCTCTGTCAGACTGTTGAAAGATGGCTTGGGGTGGGTACTTGTATAGTACTGCCACTATTGTA
ATAAAGCTAATCAAATAAAAGCACTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
ACTTCCATCTGCTCTGAAATTATGCATTTCAGGTGATTGGAGAGTCACATTCCAACTCGATGGATCCGAGTA
CTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTCTTGTATTGATGAAGTGAACAAAGTGGGCCATTGTTCAT
TTTATTGCTTGGGAAGAGCTGCTGATGAAGCAGTAAAGAGATGAGCAAGATAACAGGGAGGAGTCATTTTATGTT
TCAGATGAAGCTCAGAACAACTGGCTCATTGATGCTTGGGGCTTACATCAGGAAATACTGATCTCTCCCAGAAG
TCCCTTCAGCTCGAAAGTAAGGGATTAAACTGAACTGAACTGATGCTGGGATGAACGACACTGTCTATATTGATGACA
GTGGGAAAGGACACGTTCTTCATCACATGGAACAGTCTGCCTCCAGTATTCTCTGGGATCCAGTGGAAACA
ATAATGGAAAATTTCACAGTGGACTTCCAAATGGCTATCTCAGTATTCCAGGAACGTGCAAAGGGGGACT
TGGGCATAACATCTCAAGCCAAAGCGAACCCAGAAACATTAACATTACAGTAACCTCTGAGCAGCAAATTCTCT
GTGCGCTCAAATCAGTGAATGCTTAAAGTGAAGGACGTAACAGTTTCCCAGCCAATGATGTTTACGCGAGAA
ATTCTACAAGGATATGTAACCTGTTCTGGAGGCAATGACTGCTTCTGATGAATCACAGAAATGACATACAGAACTT
TTGGAACCTTGGATAATGGTGCAGGGCCTGATTCTCAAGAATGATGGAGTCTACTCCAGGTATTTCAGCATAT
ACAGAAAATGGCAGATATGCTTAAAGTTCAGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAATTACGGCTCCA
CTGAAATGAGCCGCTACATACCCAGGCTGGTAGTGAACGGGAAATTGAAGCAAACCCGCCAGACCTGAAATTGAT
GAGGATACTCAGACCCACTTGGAGGATTTCAGCCGAACAGCATCCGGAGGTGCTTGTGTTATCACAAGTCCCAAGC
CTTCCCTGCCTGACCAATACCCACCAAGTCAAATCACAGACCTTGTGATGCCACAGTTCATGAGGATAAGATTATTCT
ACATGGACAGCACCAGGAGATAATTGATGTTGGAAAGTTCAACGTTATATCATAAGAATAAGTGCAGTATTCTT
GATCTAAGAGACAGTTGATGATGCTTCAAGTAATACTACTGATCTGCACCAAAGGAGGCAACTCCAAGGAA
AGCTTTGCAATTAAACGAAAATATCTCAGAAGAAAATGCAACCCACATTTTATGCAATTAAAAGTATAGATAAA
AGCAATTGACATCAAACAGTATGCAACAGTAACCTGTTATCTCAAGCAATCTGATGACATTGAT
CCTACACCTACTCTACTCTACTCTACTCTGATAAAAGTCATAATTCTGGAGTTAATATTCTACCGTGGTATTG
TCTGATGGGCTGTTGAAATTGTAACCTTATTGAGTACCCATTTGAACCTTAAGAAGAAAAATCTC
AGTAGACCTAGAAGAGTTAAAAACAAAATGTAAGTAAGGATATTCTGAAATCTTAAATTCTCATCCCAT
GTGTGATCATAAAACTCATAAAATAATTAAAGATGCGGAAAAGGATACCTTGTGTTAAATAAAAACACTCATGGATA
TGTAAGGAACTGTCAAGATTTAAATTGTTGTTCTGTTATTGTTGTTATCTGAAAGAAAATAGTGTGACAAAG
ATCTTTCTACACTGATACCTGGTTGATATTGTTGATGCAACAGTTCTGAAATGATATTCTCAAATTGCACTCAA
GAAATTAAAATCATCTATCTGAGTAGTCACAAATACAAGTAAGGAGAGCAATAACACATTTGGAAAAAA
AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTYLFE
ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKG
EYIHFTPDLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRC
SAGISGRN
RVYKCQGGSCLSACRIDSTTKLYGKDCQFPDKVQTEKASIMFMQSIDS
VVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPP
VFSLLKISQRIVCLVLDKSGSMGGKDR
LNR
MNQAAKHFLLQTVENGWSVGMVHF
DSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQDGSEVLLTDGEDNTASSCIDEVKQSGAIVHF
IALGRAADEAVI
EMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDL
SQKSLQLESKG
LTLNSNAWMNDTVI
IDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTV
DATSKMAYLSIPGTAKVGTWAYNLQAKAN
PETLTITV
SR
AANSSVPPITVNAKM
N
KDVNSFPSPMIV
YAEILQGYVPVL
GANVTA
FIESQNGH
TEVLELL
DN
GA
DS
FKNDGV
Y
SRYFTAYT
TEN
GRYSL
KVR
AHGG
ANTAR
LKL
RPL
LN
RAAY
YIPG
WVN
GE
IE
AN
PP
RPE
IDE
DT
QTT
LED
FS
RT
AS
GG
AF
VVS
QV
PS
L
PL
PD
QY
PP
S
Q
IT
DL
DAT
V
HED
K
I
I
LT
WT
A
G
DN
FD
VG
KV
QRY
I
IR
IS
A
I
DL
R
DS
F
DD
AL
Q
V
N
T
D
L
S
P
K
E
A
N
K
E
S
F
A
F
K
P
E
N
I
S
E
E
N
A
T
H
I
F
I
A
I
KS
ID
K
S
N
L
T
S
K
V
S
N
I
A
Q
V
T
L
F
I
P
Q
A
N
P
D
D
I
D
P
T
P
T
P
T
P
D
K
S
H
N
S
G
V
N
I
S
T
L
V
L
S
V
I
G
S
V
V
I
V
N
F
I
L
S
T
T
I

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGTGGAAACCTGGAGTAGACTGACAGCAAAGACCGGAAAGACCATACTGCCCCGGCAGGGTGA
CAACAGGTGTCATCTTTGATCTCGTGTGCTGCCCTCCATTTCAGGAAAGACCCAAGGTAATTTCACCCCA
GAGGAGCAATGATGTAGCCACCTCCCTAACCTTCCCTCTTGAAACCCCCAGTTATGCCAGGATTACTAGAGAGTGTCA
ACTCAACCAGCAACGGCTCTTCGGCTTAACCTGTTGGAGGAGAGAACCTTGTGGGCTCGGTTCTTAGCA
GTGCTCAGAAGTGAATTGCTGAGGGTGGACCAGAAAGAAAGGAAAGGTCCCTCTGCTGTTGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAACTTGGAGATTCTCAGTCATTGCTCTGCTGCAAGATCATCTTAA
AGTAGAGAGACTGCTGTGTTGGTAACTCCAAGAGGCAAGTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCAAACGCACTGCTCTGTGCTAGCCAGGGAGGCCCTGGGGGGGGGGCTTGAGGGATGCCC
ACCGGTTCTGGACGATGGCTGATTCTGAATGATGTTGCGGGGCTGCTGCTGTTGGGATTCCGGTGGT
GTTTGCTGGTGTCTCTCTGTGCTATCTGCTGCTACATGTTGGCTGACCCAAAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGGCAACAGCCCCACGGGGAGGGTACCAAGGCCGTCCTCAGGAGTGGAGGAGCAGC
CGCAACTACGTGAGCAGCTGAAGGCCAGATGCCACAGCTCAAGGAGGAGCTGCAAGGAGGAGTGCAGCTCAGG
AATGGGCAGTACCAAGGCCAGCGATGCTGCTGGCTGGTCTGGACAGGGGCCAGAGAAAACCCAGGGCGACCTC
CTGGCCTCTGCACTCGCAGGTGGACAAGGAGGGTGAATGCTGGCTGAAGCTGGCCACAGAGTATGCAGCAGTG
CTTTGATAGCTTACTCTACAGAAGGTGACCAAGCTGGAGACTGGCTTACCCGCCACCCGAGGAGAACCTG
AGGAAGGACAAGCGGGATGAGTTGGAGGACATTGAATCAGCCTGGAGACCCCTGAACAATCTGAGAGAACAGC
CCCAATCACCGCTTACACGCCCTCTGATTTCATAGAAGGGATCTACGAACAGAAAGGGACAAGGGACATTGAT
GAGCTCACCTCAAGGGGACACACAGGAAATTCAACAGCTCATCTTATTCGACCTTCAGCCCCATCATGAAA
GTGAAAATGAAAAGCTAACATGGCCAACACGCTTATCAATGTTATCGTGTCTAGAAAAGGGTGGACAAGTTC
CGGCACTGATGCAAACTTCAAGGGAGATGTCATTGACAGGATGGAGACTTCACACTGTTGTTACTTGGG
AAAGAAGAAAATGAAGTCAAAGGAAACTGAAAACACTTCCAAAGCTGCCACTTCAGGAACTTACCTCATC
CAGCTGAATGGAGAATTTCGGGGAAAGGAACTTGATGTTGGAGCCCGCTTCTGAAGGGAAACCACTCCTCTC
TTTTCTGTGATGGACATCTACTCACATGTAATTCTCAATACGTTGAGGCTGAATACACAGCCAGGGAAAG
GTATTTTATCCAGTTCTTCAGTCAGTACAATCTGGCTAATAATACGGCCACCATGATGCACTCCCTCCCTGGAA
CAGCAGCTGGTCAAAAGAAGGAACTGGATTGGAGAGACTTTGGATTTGGATGACGTGTCACTGTTGTTACTTGGG
TTCATCAATATAGTGGGTTGATCTGGACATCAAAGCTGGGGGGAGAGGATGTCACCTTATCGCAAGTATCTC
CACAGCAACTCATAGTGGTACGGACCCCTGTCGGAGGACTCTTCCACCTCTGGCATGAGAAGGCCGATGGACGAG
CTGACCCCCGGAGCTAACAGATGTGATGCACTGCAAGGCCATGAACGGCAGCTCCAGGGCAGCTGGGATGCTG
GTGTTCAAGCAGAGATAGGGCTCACCTCGCAAACAGAAAAGACAAGTAGACAAAAGACATGAACTCCAGA
GAAGGATTGTGGAGACATTTTCTTCCCTTGCACATTACTGAAAGTGGCTGCAACAGAGAAAAGACTCCATAAA
GGACGACAAAAGAATTGGACTGATGGGTCAGAGATGAGAAAGCCTCGATTCTCTGTTGGCTTTTACAACAGA
AATCAAATCTCCGTTGCTGCAAAGTAACCCAGTTCAACCCCTGTGAAGTGTCTGACAAAGGAGAATGCTTGTG
AGATTATAAGCTAATGGTGGAGGAGTTGATGGTTTACAATACACTGAGAGACTGTTGTTGCTCATG
AATATTCTGATTTAAAGAGCAGTTGAGATGCAATTGAGACATTGAGCATGAAAGGAGCATATTCTCTCATATGAATG
GCCTATCAGCAGGGCTCTAGTTCTAGGAATGCTAAATATCAGAAGGAGGGAGAGATAGGCTTATTATGATACT
AGTGAGTACATTAAGTAAAATGGACCAAGAAAAGAAAAGACATAATATGTCATATTCTCCCAAGAT
TAACCAAAAATACTGCTTATCTTGGTTGCTCTTAACTGTCCTGTTTCTTCTTATTTAAAGTGCAC
TTTTCTCTGTGAGTTATGCTGCTTATTAATTACCACTTCAAGCCTTCAAGAGAGCAGGACAAAGTGGCTAC
ATTTTATATTAAAGAGAATCTGAGATGCAATTGAGACATTGAGCATGTTCAAGGACATCAAATTGATGCCATAT
CCAAGGACATGCCAATGCTGATTCTCAGGCACTGAATGTCAGGCAATTGAGACATAGGGAGGAAATGTTGACT
AATACAGAGTACAGATACTTCTGTAAGAGTATTTCGAAGAGGAGGAGCACTGAACACTGGAGGAAAAGAAAATGAC
ACTTCTGTTTACAGAAAAGGAAACTCATTCAAGACTGGTGAATCTGATGTCACCTAAAAGTCAGAAACCACTT
CTCCTCAGAAGTAGGGACCCCTTCTACCTGTTAAATAACCAAGTATACCGTGTGAACCAAACAAATCTCTTTC
AAAACAGGGTGTCTCTGGCTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATTGT
GAAAGATCAATCCATCTGCCAGAACTGAGTGGGATGGAATTGCTACATGTTATCCACCCAGGGCAAGTGGAG
TAAGTGAATTATTTAAAGCAGTTCAACTCAATCAGGAGATGCTCTGAAATTGCAATTG
CAAACATTTTAAAGAATACAGTTAACATAGAGTGGTTCTCATTGATGAAAATTATTAGCCAGCAG
ATGCATGAGCTAATTATCTCTTGTGAGTCCTGCTTCTGCTACAGTAAACTCATTGTTAAAGCTTCAAGAAC
ATTCAAGGCTGTTGGTGTGTTAAAGATGCAATTGATTTGACTGGTAGTTATGAAATTAAATTAAACACAGG
CCATGAATGGAAGGTTGATGCAAGCTAATAAAATATGATTGTTGGATATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESALETLNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNELNMAN
TLINIVPLAKRVDKFRQFMQNFRMCMIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKVFYPVLFQSQY
NPGIITYGHHDAVPPLQQQLVIKKETGFWRDFGFGMTQYRSDFINIGGFDLGIKGWGGEDVHLYR
KYLHSNLIVVRTPVRGLFHILWHEKRCMDELTPEQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGAGTTGTCTGGGATCCA
GAAACCATGATAACCTACTGAACACCGAATCCCCTGGAAGGCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCGCTGTCTAGTCCTAGTCCTAAATTCCAGTCCCTGCACCCCTC
CTGGGACACTATGTTCTCCGCCCTCTGCTGGAGGTGATTTGGATCCTGGCTGCAGATGGG
GTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATTGCCAGCCTCTTACCCGTAGTGT
GGAAACATGCCAGTCGCCATCGATATTCAAGACAGACAGTGTGACATTGACCTGATTTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTCAACTCTCTGCCCTTACCCGTATCTGGTGGACTCCCCGAAAATATGTAGCT
GCCAGCTCCACCTGCACTGGGTCAAGAAAGGATCCCCAGGGGGTCAGAACACCAGATCAACAG
TGAAGCCACATTGCAGAGCTCCACATTGTACATTGACTCTGATTCCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGCCCTCAGGGCTGGCTGCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCATTGCATGAAGTCAGGCATAAAGATCAGAACACTC
AGTGCCTCCCTCAACCTAACAGAGAGCTGCTCCCCAAACAGCTGGGAGTACTTCCGCTACAATG
GCTCGCTCACAAACTCCCCCTGCTACCAGAGTGTGCTTGGACAGTTTTATAGAAGGTCCAG
ATTTCAATGGAACAGCTGGAAAGCTTCAGGGACATTGTTCTCCACAGAACAGGGAGCCCTCAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGCATGGTCTTGCTTCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTGGTGGC
TGTCTGCTCCTCTGGCTGTTATTCATTGCTAGAAAGATTCCGAAGAACAGGGCTGGAAAA
CCGAAAGAGTGTGGTCTCACCTCAGCACAGCCACGACTGAGGCATAAATTCCCTCTCAGATAC
CATGGATGTGGATGACTTCCCTCATGCCTATCAGGAAGCCTCTAAATGGGTGAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCTCTCCCTGGACATCTTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAACGACTGCAGAGCCTCAGCCTCTCCAAACATGTAGGAGGAA
ATGAGGAAATCGCTGTTGTTAATGCAGAGANCAAACCTGTGTTAGTGCAGGGGAAGTTGGG
ATATAACCCAAAGTCTTACCCCCCTCACTTTATGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGATATTTGATCAATATAATTGGAAATTAAAG
TTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECNNNAQSPIIDIQTDSVTFPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCCCTGGGACGGCAGTCCCTGTGTC
TCTGGGGTTGCCTAACCTGAAACATCACCTCTTATCCATCAACATGAAGAATGTCCTACA
ATGGACTCCACCAGAGGGTCTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTCATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGTCTGACAGCTCC
AGAGAAGTGAAGAGAAATCCAGAAGACCTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTCTGTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGACCAACCAC
ACGCTGGTGCACCTGGCTGGAGCGAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTGCCGTGTCAGCCTCTGAGAAGCAGTGTGCCAGGACTTGAAGATCAATCAT
CAGAGTTCAAGGCTAAATCATCTCTGGTATGTTGCCATATCTATTACCGTGTCTTT
TCTGTGATGGGCTATTCCATCACGATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
GATTTGATTTATGAAATGAATTGACAAAGATTCTTGTGCCTGCTGAAAAAATCGTGATTA
ACTTTATCACCCCTCAATATCTGGATGATTCTAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGGCCCCCTCAGGA
GGAAGAGGGAGGTGAAACATTAGGGTATGCTTCGCAATTGATGAAATTGGACTCTGAAG
AAAACACCGAAGGTACTTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTGTGCGGGGCCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGTGTCCACACAAGGAACATTGGAGTCGCAGGCAGCGTGGCAG
TCTTGGCCCCCAAACGTTACAGTACTCATACACCCCTCAGCTCAAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGGCCATCGACGACCCCTGGTCAGTGGGA
TCCCCAAACTGGCAGGCTGTGTATTCCCTCGCTGTCCAGCTCGACCAGGATTAGAGGGCTGCG
AGCCTCTGAGGGGGATGGGCTGGAGAGGGAGGTCTTCTATCTAGACTCTATGAGGAGGCCGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCCAACACTCCCTTTGCCTTGTGCAAACAAAGTGAG
TCACCCCTTGATCCCAGCCATAAAAGTACCTGGGATGAAAGAAGTTTTCCAGTTGTCACTG
CTGTGAGAATTACTTATTCTTCTTCTATTCTCATAGCACGTGTGATTGGTCATGCATGTA
GGTCTCTAACATGATGGTGGCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
AGCAGTCATAAAATGTTGCCAGACTGGGTGCAGAATTATTACAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSISSQIGPPEVALTTDEKSISVVLTapekwkrnpedlpvsmqqiy
SNLKYNVSLNTKSNRTWSQCVTNHTLVLWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGYNEFDKRFVPAEK
IVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEVKHTMLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLQEEVSTQGTLLESQA
ALAVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDSLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQmen

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTGACAC
CTGGGAAGATGGCGGGCCGTGGACCTTCACCCCTCTCTGTGGTTGCTGGCAGCCACCTGATC
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCGGGAGGCATCCCTGTGCTGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
ATCTGGCTGAAGGTCACTCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCTCGGCCAATGACCA
GGAGCTGCTAGTCAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAAGGCCACCATCCGCATGGACACCAGTGAAGTGGC
CCCACCCGCCTGGCCTCAGTGACTIONTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTGTGA
TAAGCTCCTCCTGGTGAACGCCCTAGCTAACAGCTGCAGGTCAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGAAAACCAGCTGTGCTCCGTGATCGAGGCTCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGGTACACCATTCAAGCTTACCTGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTCAATAACTCTGCACTTCCCTGACAATGCCACCCCTGGACAACATCCCG
TTCAGCCTCATCGTAGTCAGGACGTGGTAAAGCTGCAGTGGCTGCTGTGCTCTCCAGAAGA
ATTCAATGGCCTGTTGGACTCTGTGCTTCTGAGAGTGCCATGGCTGAAGTCAGCATGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTAAGATCTTAAGTCAAGC
ACTCCCGAGTTTTTATAGACCAAGGCCATGCCAAGGTGGCCAACTGATCGTGTGGAAGTGT
TCCCTCCAGTGAAGCCCTCCGCCCTTGTGCTTCAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTAACTCAACTTGAATAACATCAGCTGTGATCGGATCCAGCTG
ATGAACCTGGGATTGGCTGGTCCAACCTGATGTTGCTGAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAAATGGCAAATTAAAGATCTGGGTCCCAGTGTGATTGGTGAAGG
CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTGTGCTTACTCCAGCTCC
TTGTGGAAACCCAGCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATAGGGAAAGGCTGG
GTCCCAAGCTGGAGTATGGGTGTGAGCTATAGACCATCCCTCTGCAATCAATAACACTTG
CCTGTGAAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPVKVIKEKLTQELKDHNATSILQQPLLSAMREK
PAGGIPVLGSLVNTVLKHIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPILDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPESAHRLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLGIASSEAQFYT
KGDQLILNLNNISSLRQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTCTGGGCTTTGGGACACTGGTGCAT
GCTGCTCCCAGCTGGAAAACAAGTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTCT
CCAAGGGCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCTCTGGGCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAT
CTCCTCCCTGGCCTGCATTATCTGTGGTGGCATGAGATGCACAGTCTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTCATCCTGGAGGCCTGGGATTC
ATTCCCTGTTGCCTGGAATCTTATGGGATCCTACGGGACTTCTACTCACCACGTGGCCTGACAG
CATGAAATTGAGATTGGAGAGGCTCTTACTTGGGATTATTCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTCTGCTCATCCAGAGAAATCGCTCAACTACTACGATGCCTAC
CAAGCCAACCCTTGCCACAAGGAGCTCCAAGGCCGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGTAAGAACCAGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTAAAAACAGTGGACAGCACCCGAGGGCACAGGTGAGGGACACTACCAACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTGGCATTGGATTGAGCAAAGGCAGAAATGGG
GCTAGTGTAACAGCATGCAGGTGAATTGCCAAGGATGCTGCCATGCCAGCTTCTGTTCC
TCACCTTGCTGCCCTGCCCTAAGCCCCAACCTCAACTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGCCCTGGTTACCTGGACTCCATCCCCAAACCCACTAATCACA
TCCCACTGACTGCCCTGTGATCAAAGACCCCTCTGGCTGAGGTTGGCTTAGCTCATT
GCTGGGATGGGAAGGAGAAGCAGTGGCTTGTGGCATTGCTTAACCTACTCTCAAGCTTC
CCTCCAAAGAAACTGATTGCCCTGGAACCTCCATCCACTCTGTTTATGACTCACAGTGTCCA
GACTAATTGTCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGACATTAAAAAAATA

FIGURE 80

MASLGLQLVGYILGLLGLLGTIVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSAAISLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGLRDFYSPLVPSMKFEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCACCGCGTCCGCCTCTCCCTCTGCTGGACCTCCCTCGTCTCCATCTCTCCCTCTTC
CCCGCGTCTCTTCCACCTTCTCTTCCACCTTAGACCTCCCTGCCCTCCCT
GCCACCGCGTCTCCCTGGCCCTCTCGACCCGCTCTAGCAGACCTCCCTGGGCTGTGG
GTTGATCTGTGGCCCTGTGCCCTCGTCTTCTGCTCCCTCCGACTCCGCTCCCG
ACCAGCGGCCTGACCTGGGAAAGGAGGTGGTCCCGAGGTGAGGGCTCTCCTGCTGGGA
CTCGCGCTGCTCTGGTCCCCCTGGACTCCACGCTCGAGCCGCCAGACATGTTCTGCCTTT
CCATGGGAAGAGATACTCCCCCGCGAGAGCTGGCACCCCTACTTGAGGCCACAAGGCCTGATGT
ACTGCCTGCCTGTACCTGCTCAGAGGGGCCATGTGAGTTACCGCCTCACTGTCCGCC
GTCCACTGCCCTAGCCTGTGACGGAGCCACAGCAATGCTGTCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCCCGCCACCAAAGTCTGCCAGCACACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCATGAGCTGTTCCCTCCGCCATGCCAACAGTGTGTCCCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCCCTCACAAACCTGCCGAACCAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCTGCTGCCAAGCCTGCAAAGATGAGGAAGTGAAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCCAGTGTGCTGGAGAAAG
AGAGGCCCGGGCACCCAGCCCCACTGCCCTAGGCCCTCTGAGCTTCATCCCTGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAGATCGCTGAAGGAGAAACATAAGAAAGCT
GTGTGCATGGCGGGAAAGACGTACTCCCACGGGAGGTGTGGCACCCGCCCTCCGTGCCCTCGGC
CCCTGCCCTGCATCTATGCACCTGTGAGGATGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCAGAAAGTGGCTGGAAAGTGTGCAAGATTGCCAGAGG
ACAAAGCAGACCTGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCGGTC
CTCGTCCACACATCGGTATCCCCAAGGCCAGACAACCTGCGTCGTTGCCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCAAGGCCACACAGCCAGAACATCTCCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTCCAGAAAGAGGCCACAGCACCTCCGACTGCTCGTGGCCCCACGAAGGTCACT
GGAACGTCTCCTAGCCCAGACCCCTGGAGCTGAAGGTACGGCCAGTCCAGACAAAGTGAACCAAG
ACATAACAAAGACTAAACAGTTGCAGATATGAGCTGTATAATTGTTATTATATAATAAAA
TAAGAAGTTGCATTACCCCTCAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELP
PSRLPNQCVLCSCTEGQIYCGLTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTY
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEEETEAQRGEVPGPRPH
SQNLPLSDQESQEARNLPERGTALPTARWPPRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTCTCTAATCCATCCGTACCCCTCCTGTCA
TCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTGGTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGGCCAGACAAGCCTGTCAGGGCTTGGTGGGGAG
GACGCAGCATTCTCCTGTTCCCTCTAAGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTCAGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAATGGTGAAGGATTCTATTGCGGAGGGCGCATCTCTGAGGCTGAAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTCCCAGTCTTACTACCAGAAGGCCATCTGGAGCT
ACAGGTGTCAGCACTGGCTCAGTCCCTCATTTCCATCAGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCCTGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCACAAGGACAGGATTGTC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTGATGTGGAGATCTCTGACCGTCCAAGAGAA
CGCCGGAGCATATCCTGTTCCATGCCGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTGAGCCTATATCGTGGCACCTGGCTACAAAGTACTGGAATACTCTGCTGTGGCTA
TTTTTGGCATTGTTGACTGAAGATTCTCTCCAAATTCCAGTGGAAATCCAGGCCAACTGGACT
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGAAACACGCACTGGAGGTGACTCTGGATCCAG
AGACGGCTCACCGAAGCTCTGCGTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGTGGCGCGTGGAGTGTGCCCCGATGATGGACAGGAGGA
AGGAGTACGTGACTTGTCTCCGATCATGGTACTGGTCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAATCCCCGTTTATCAGCGCTTCCCCAGGACCCACCTACAAAAATAGGGTCTTCTGGACTA
TGAGTGTGGGACCATCTCCTCTCAACATAATGACCGACTCCCTATTATACCCGATGTCGGTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCGTCTATAATGAGAAAATGGAACCTCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCCTTGGCAAAGGGCCTCTGCAATCCCAGAGACAAACAG
TGAGTCCTCCTCACAGCAACCACGCCCTCCTCCCCAGGGTGAATGTAGGATGAATCACATCCCACAT
TCTTCTTAGGGATATAAGGTCTCTCCAGATCAAAGTCCCGCAGCAGGCCAAGGTGGCTTCCA
GATGAAGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTAGCTGGCTGCCCTGAGCTGGGAGGGAAAGG
CTGACATTACATTAGTTGCTCTCACTCCATCTGGCTAAGTGTAGCTTGAATACCACCTCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCTAATGGTTGTTCAATTACACTTCAGTAAAAAA

FIGURE 84

MALMLSIVSLLKLGSQWQVFGPDKPVQALVGEDAASFCLSPKTNAEAMEVRFFRGQFSSVH
LYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFFRPTAKWKGPGQGDLSTDRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAEQELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVVASQSFQAGKHYWEVDGGHNKRWRVGVCVRDDVDRKEYVTLSPDHGYWVLRLNGEHLYFT
LNPRFISVFPTPPTKIGVFLDYECGTISFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGGGCCCTGGCACCTCAACCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTCTCCTACCCCTCGCATGGCTGGATT
TACCCCTGGCCAGTAGTTCATGGCTACTGGTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTG
GGGACCCACATACCAAGAATTGACCCCTGAGCATCAGAGATGCCAGAAGAAGTGTGCGGGGAGA
TACTTCTTCGTATGGAGAAAGGAAGTATAAAATGAAATTATAAACATACCGGCTCTGTGAA
TGTGACAGCCTGACCCACAGGCCAACATCCTCATCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGCCTGTGAGCAGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCACCACCCGCTCCTCGGTGCTCACCTCATCCCACA
GCCCCAGGACCATGGCACAGCCTCACCTGTCAAGGTGACCTTCCCTGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGCTCACCGCCTCAGAACATTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTC
GCGCCTGGTCTGTGAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCCTGTGCCCCCTCACAGCCCTAAACCCGGGGTGCTGGAGCTGCCCTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCCTCTGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGAGCTGGAG
CCACAGCCCTGGTCTCCTGTCCCTGCGTCATCTCGTTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCTGGGAGATAACGGGCTAGAGGATGCAAACGCTGTCAGGGGTT
AGCCTCTCAGGGGCCCTGACTGAACCTGGCAGAAGACAGTCCCCAGACCAGCCTCCCCAG
CTTCTGCCGCTCTCAGTGGGAAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGT
AAGCCTTGGACTCGCGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTGTAGAATTAAACAGCCCTAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAGAACACAGGCTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTTATTTTTAACTAAAGACAGACAAATTCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTQEGLCVHPCFSYPSHGWYIPGPVVFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLISRARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNTALTHRPNILIPGTLESQCPQNLTCSPWACEQGTPPMISWIGTSVSPLDPSTRS
SVLTLIPQPQDHGTSLTCQVTFFGASVTTNKTvhlnvSYPPQNLTMTVFQGDGTVSTVLNGSSL
SLPEGQSLRLVCAVDADSNPPARLSLSWRGLTLCPSPNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVRGSASQGPLTEPWAEDSPPDQPPPASARSSVGEGELOQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGAGTGGAGGGAGGGAGTGAAGGAGCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTGCCCAGAAGCTGCAAGGAATCAAAGACGAATGTCCTAGTGCATTG
TGGCCTGTATTTCTCGCACTGAGAATGGTGTATCTACCAAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGT
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCAACTGGC
CAAATACAACACCTTGGATCTGCAGAGGCCACGAGCGATGACTACAAGAACCCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTCCAGACACTGGGACATAATCT
GTTGGCATCTACCAAGAAATATCCAGTGAATATGGAGAAGGAAAGTGTGACTGACAACGCC
CGGTGATCCCTGTGGTCTATGATTTGGCAGGCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGAATTCACTGCGGATTGTTCACTCAGGGTATTTAATAACGAGAGAGCAGCAA
CGCCTTGTGTGCTGGAATGAGGGTACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTCTGGTTTGATTGGAGTGGATAT
GGAACTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAACCCAGACCTCTCCCAACCAGAGATCCAAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENMRGKCTVGRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIKVYDFGDAQKTASYSPYGOREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGC~~GGGG~~AGACTTCAGGAGTCGCTGTCTGAAC~~TTCCAGC~~CTCAGAGAC
CGCCGCCCTGTCCCCGAGGGCCATGGGCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTATTCTGGTCCC~~GGG~~ACAGCAACATA~~CAGG~~CCTG
CCTGCCTCTCACGTTCACCCCCGAGGGAGTATGACAAGCAGGACATT~~CAG~~CTGGTGGCCGCGCTCT
CTGTCACCCCTGGGCCTCTTGCAGTGGAGCTGGCGTTCCCTCTCAGGAGTCTCCATGTTAAC
AGCACCCAGAGCCTCATCTCCATTGGGCTCACTGTAGTGCA~~TCCG~~TGGCCCTGTCCTTCTTCAT
ATT~~CG~~AGCGTTGGGAGTGC~~ACT~~ACGTATTGGTACATTTTGCTTC~~TG~~CAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTATT~~CG~~TACCGT~~TTGG~~CTGAAAAAGAAACCCCTGATTACCTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCGCTCGTATT~~CC~~TGGAAGAAGGAAG
GCATAGGCTCGGTTTCCCTCGAAACTGCTTCTGCTGGAGGATATGTGTTGAAATAATTACG
TCTTGAGTCTGGATTATCCGATTGTATTAGTGCTTGTAAATAATGTTTAGTAACA
TTAAGACTTATATACAGTTAGGGACAATTAAAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFC SALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCGAAAAGAGAAGGGAGAGCGAGGGACGAGAGCGGAGGAAGATGCACTGAC
TCGCTGCTCGTGTCTGGTGCAGGGTAGCCTCTATCTGGTATCTGTGGCAGGATGATG
GTCCTCCCGCTCAGAGGACCTGAGCGTGTGACCACGGGCCAGCCCCGGCCCCGGTGCCT
CGGAAGGGGCCACATCTCACCTAACGCCCCATGGCAATTCCACTCTCTAGGGCTGCT
GGCCCCGCTGGGGAGGCTGGGCATTCTGGCAGCCCCCAACCGCCGAACCACAGCCCC
CACCTCAGCCAAGGTGAAGAAAATCTTGCTGGGCACCTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTGTCACAGGAAAGATTGTGGACCATGGCAATGGACCTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCATCAGCCTCGTGCCTCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTCATCGAACGCCCTCAAATCTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGCCGCGACCTCGTTGCACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGGCCACTGGAGCTGCTCCAGCCCTCAAAGTCGTCTGTG
TCTACATCGCTTCTACAGCAGGACTATCGGCTGGTCAGAAGGTGTGCCAGATTACAACCTAC
CATAGTGTACCCCCACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAGGACAGGCCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAGGGGTTGGCCTC
AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCAAGTGTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGCTCTGTGCAAGCCTCACAGGCTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCCCTGTGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAC
CCTGGTTCTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGAGATTCTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCCAGCCCTGCCTAGTGGCGCCCTGAGCCCTGTGCTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCCCTGGGTCTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCAAATCCCTTTCTGCCAGTACTCCCCCTGTACCAACCCATTGCTGATG
GCACACCCATCCTTAAGCTAACGACAGGACGATTGTGGTCCCTCCCACACTAACGGCCACAGCCCAC
CGCGTGTGTGTGTCCTCTCCACCCCAACCCCTGCTGGCTCCTCTGGAGCATCCATGTCCCG
GAGAGGGTCCCTAACAGTCAGCCTCACCTGTCAGACCGGGTTCTCCGGATCTGGATGGCG
CGCCCTCTCAGCAGCGGGCACGGTGGGGGGGGCGAGACATGTGCTGGATCTGTC
TGTGTGTCTGTGTGGTGGGGGGAGGGAGGGAAAGTCTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGAGCAGGAATAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNNTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVVCVYIAFYSTDYRLVQKCPDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGC~~GGCC~~GTGTTCTCGGCTGCGCCTTCATTGCCTTCGGCCTGCGCTGCC
CTTTATGTCTTCACCATGCCATCGAGCCGTTGCGTATCATCTTCCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATTCGTCCTTGTGTTGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGGAGCGTTGCTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTATGCGACTGCTGGCCTATGTTCTGGCTTGGGCTTGGAAATCATGA
GTGGAGTATTTCTTGTGAATACCCATCTGACTCCTGGGCCAGGCACAGTGGCATTCAT
GGAGATTCTCCTCAATTCTCCTTATTCA~~GCTT~~CATGACGCTGGCATTATCTGCTGCATGT
ATTCTGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTCA~~GCCC~~AGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTATAATCCTGGTGTCA~~GGG~~ACCTGGCATTCTAGCTGCGGAGGCAGCTGCCG
AACCTGAAACTCTGCCTGCTCTGCCAAGACAAGAAACTTCTTACAACCAGCGCTCCAGAT
AAACCTCAGGAACCAGCACTCCAAACCGCAGACTACATCTTAGAGGAAGCACAACTGTGCCT
TTTCTGAAAATCCCTTTCTGGTGAATTGAGAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMLLAVSGLGFGIMSGV
FSFVNNTLSDSLGPVTGVIHGDSPQFFLYSAFMTLVIILLHVFVWGIVFFDGCEKKWGILLVLLT
HLLVSAQTFISSLYYGINLASAFIILVLMGTWAAGGSCRSLKLCLLCQDKNFLLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTGTCTTAGGATCAAC
TCGGTCATTACCACAGCTAAACCTGCTTGGACTCCCTCCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAACCAACAGCAGTAAATCAGGTCTTCCTTCTTAAGTCTGATACCATTAAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCCTGGGAGGGTTGAATGTACAAACAGCAACTGCACCCACATGTGTAC
CAATTTGTCACACAACTGGAGGCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTGTTCCCGGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG
GAACCCCAGCAGGCCCTCCAACTCCCAGTGGCACAGATGACGACTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTCAAATTTTCAACTAAGCTGCCTCGAATTGGTGATACTGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT
TACCTGAAAATATTCTTGAAATTCAGAAAATATGTTCTATGTAGAGAATCCAACTTTAAAAAA
CAATAATTCAATGGATAAAATCTGCTTGAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTGAAAACTGGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSSLIPLTQM
LTLPDHLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCCTGCCCTGCCCAACCCAGGCCAGGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTGCTGGCAGCTGGACCAAGGGAGCAGTCTGGCGCTGGAGGGCTGTCAGCAT**GATG**
GTCCCTGCCCTGGCTGTGGCTGTTGTCTCCGCCCCAAGGCTCTCCCCAAGGCCAGCTGC
AGAGCTGTCTGTGGAAGTTCAGAAAATATGGTGGAAATTCCCTTATACCTGACAGTTG
CGCTGCCCGTGAAGGCCAGATCGTGTGTCAGGGACTCAGGCAAGGCAACTGAG
GGCCCATTTGCTATGGATCCAGATTCTGGCTCCTGCTGGTGAACAGGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACCACTACAGGTCACCCGGAGATGCAAGGATGGACATGTTGTTGGGTCAC
AGCCTGTGTTGTGACGTGAAGGATGAGAATGACCAGGTGCCCAATTCTCTCAAGCCATCTAC
AGAGCTCGGCTGAGCCGGGTACCAAGGCCAGGGCATCCCTTCTGAGGCTTCAGACCG
GGATGAGCCAGGCACAGCCAAGCTGGATTCGATTCCACATCTGAGCCAGGCTCCAGGCCAGC
CTTCCCCAGACATGTTCCAGCTGGAGCCTGGCTGGGGCTCTGGCCCTCAGCCCAAGGGAGC
ACCAGCCTGACCACGCCCTGGAGAGGACCTACCAGTGTGGTACAGGTCAGGACATGGGTGA
CCAGGCCTCAGGCCACCAGGCCAGTGCACCGTGGAGCTCCATAGAGACACCTGGGTGT
CCCTAGAGCCTATCACCTGGCAGAGAATCTCAAAGTCTTACCCGACCATGGCCAGGTA
CACTGGAGTGGGGGTGATGTCAGTACATCACCTGGAGGCCATCCCCGGGACCCCTGAGTGA
TGCAGAGGAAACCTTACGTGACCAGAGAGCTGGACAGAGAAGGCCAGGCTGAGTACCTGCTCC
AGGTGCGGGCTCAGAATTCCATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACTGTCGTC
ATGGATGAGAATGACAACGTGCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGAAGTGAAGTGTCACTGGAGGATGAGATGCCCGGCTCCCCCA
ATTCCACGTTGTGATCAGCTCTGGACCCCTGAGCTGAGGATGGGGTAGAGGGGAGACCTTC
CAGGTGGACCCCACTTCAGGCACTGTGACGGCTGGGGTCTCCACTCCAGCAGGCCAGAACAT
CTGCTCTGGTGTGGCCATGGACCTGGCAGGGCAGAGGGTGGCTCAGCAGCACGTGTAAG
TCGAAGTCGCACTCACAGATATCAATGATCACGCCCCCTGAGTTCATCACTCCCCAGATTGGGCT
ATAAGCCCTCCCTGAGGATGTGGAGCCGGACTCTGGTGGCCATGTAACAGCATTGATGTCGA
CCTCGAGCCCCGCTTCCGCTCATGGATTTCGCCATTGAGAGGGGAGACACAGAAGGGACTTTG
GCCTGGATTGGGAGCCAGACTCTGGCATGTTAGACTCAGACTGCAAGAACCTCAGTTATGAG
GCAGCTCCAAGTCACTGGGGTGGGGTGGTGGTGGCAGAGTGTGGCAAGGCTGGGGCCAGGCC
AGGGCCCTGGAGGCCACCGCCACGGTCACTGTGCTAGTGGAGAGTGTGGCAAGGCTGGGG
ACCAGGAGAGCTACAGGGCCAGTGTCCCCATCAGTGCCTGGAGGGCCAGGATACAGCCTGACT
CAGCCCCCTCCGACCCCATCAGCGAACCCCTCAGGTTCTCCCTAGTCAGTCACTGAC
CTGCAAGGAGGCTGG
ACACACTACAGGTGTTGTGGAGGGCCAGGATACAGCCTGACTCTGGCCCTGTGCCCTCCAA
TACCTCTGACACCCCGCCAAGACCATGGCTTGTGAGTGGACCCAGCAAGGACCCGATCT
GGCCAGTGGGAGCCAGGCTCCCTACAGCTTACGCTTACGCTTGGTCCAAACCCACGGTCA
GGCTCCAGACTCTCAATGGTCCCATGCCTACCTCACCTGGCCCTGCATTGGGGAGGCCAGT
GAACACATAATCCCCGTGGTGGTCAAGGACAGGACCCGGATCAACCCAGCAGACAGCGT
CGTGTGTCGCTGCACTGGAGGGGAGGGCAGTGCATGCGCAAGTGGGGCGCATGAAGGG
CGAAGCTGTCGGCAGTGGCATCTTGTAGGCACCCCTGGTAGCAATAGGAATCTCCTCATCCTC
ATTTTCAACCCACTGGACCATGTCAGGAAGAAGGACCCGGATCAACCCAGCAGACAGCGT
GAAGGGCAACTGTCTG**GA**ATGGCCCAAGGCAAGCTAGCTGGAGCTGGGCTCTGGCTCCATCTGAG
TCCCTGGGAGAGAGGCCAGCACCCAAAGATCCAGCAGGGACAGGACAGAGTGAAGGCCCTCCA
TCTGCCCTGGGTGAGGGCACCACATCACCAGGCACTGTCAGGGTGTGGGCCATACCAACTT
TATGGACTGCCATGGGAGTGTGCTTCAATGTCAGGGTGTGGGCCATATAAAGGCCAGAGA
CTGGGCTGGGCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLREGAEQQIVLSGDGKAT
EGPFAMDPDSGFLLVTRALDREEQAELYQLQVTLEMQDGHLWGPQPVLHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAELYLLQVRAQNSHGEDYAAPLEHVL
VMDENDNVPICPPRDPPTVSIPELSPPGTEVRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGVTLGVPLRAGQNLILLVLAGAEGGSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTAIDADLEPAFRIMDFAIERGDTEGTFGDWEPSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPAGATATVTVLVERVMPPPQLDQESYEASVPISAPAGSFL
IQPSDPISRTLRFSLVNDSEGWLICIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTLPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTLALHWEP
REHIIIPVVVSHNAQMWWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGT
LIFTHTMSRKKDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCACGCTCTGGAAAGCACCAGCCTTA
TCTCTCACCTCAAGTCCCCCTTCTCAAGAATCCTCTGTTCTTGCCTCTAAAGTCTGGTAC
ATCTAGGACCCAGGCATCTGCTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTATGTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACAGCCACCATCTCAGGGTCCAGCGTACCTCCAATGGGTCA
GCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCCACACCTCCAGTGGGCCAGCACAGTACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGCTAGCACAGCCAC
CAACTCTGACTCCAGCACAACCTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCCACACTGCCACCAACTCTGAG
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACACCTCCAGTGGGCCGGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCC
AGTGGGCCACACAGCCACCAACTCTGAGTCCAGTACGACACTCTCCAGTGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGTCAAGCACAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCAC
GCCACCAACTCTGACTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCTCCAGTGGGCCA
ACACAGCCACCAACTCTGGTCCAGTGTGACCTCTGCAGGCTCTGAAACAGCAGCTGACTGGA
ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAGTGAGGCAAAGCCTGGTGGTCCCTGGT
GCCGTGGAAATCTTCTCATCACCCCTGGTCTGGTTGGCGGCCGTGGCTCTTGCTGGC
TCTTCTCTGTGTGAGAAACAGCCTGTCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
CATGGCTCAACCATGGCTTGGTCCAGGCCCTGGAGGAATCATGGAGCCCCCACAGGCCAG
GTGGAGCTTAACCTGGTCTGGAGGAGACCAGTATCATCGATGCCATGGAGATGAGCAGGGAGGA
ACAGCAGGGCCCTTGAGCAGCCCCGAAGCAAGTGCAGCATTCTCAGGAAGGAAGAGACCTGGCA
CCCAAGACCTGGTTCCCTTCAATTCCAGGAGACCCCTCCAGTTGTTGAGATCTGAA
AATCTGAAGAAGGTATTCCCTCACCTTCTGCCTTACAGACACTGGAAAGAGAATACTATAT
TGCTCATTAGCTAAGAAATAACATCTCATCTAACACACAGCACAAAGAGAAGCTGTGCTTG
CCCCGGGGTGGTATCTAGCTGAGATGAACTCAGTTAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAATCTCCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTNGAGTATNSES
STTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTTSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVA
AVGLFAGLFFCVRNLSLRNTFNTAVYHPHGLNHLGPGPGENHGAPHRPRWSPNWFWRPVPSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCCTCCCGTTACGGGATGAATTAAACGGCGGTTCCGCACGGAGGTGTGACCCCTA
CGGAGCCCCAGCTTGCACCGCACCCACTCGCGTCGCGCGTGCCTGCTGTACAGGTG
GGAGGCTGAACTATCAGGCTAAAAACAGAGTGGTACTCTCTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAGCTTGGAAATCAT
GGTGTATGGAAAGGGATTACTTATACGTACTCTGTTGGGAAGCTTTTGGAAAGCATT
TCATGCTGAGTCCCTTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTTGGCAACATGGCTACCCCTACCTGGCATTATTGGAGACCAGTTGGTAAAGT
GATTATAACTGGGATGCATTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCAGCGACAA
GAATGGACTGGATGTTCTGTGAAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCAAGGAAATGGAAGGATGACAAGAGCCATTGCAAGACATGATTGATTACTTTGT
ATATTCAACGAACCACCTCAACTCCTCATATCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTACTTGTGGTAGACCGTCTAAGAGAAGGTAAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGGTATCCTCACACATTCCCAATCAGAGAACCTCCCTCCAAGGAGACTT
CCCAGGAAATCCACTTCACGTCACCGGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCACAAACGGTGGAGAGAAAGAAGAGAGAGGCTGCGTTCTTCTATCAAG
GGGAGAAGAATTTTATTTACCGGACAGAGTGTCAATTCCACCTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCTGTTAGCCCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTTAAGGGTATTATAATCACCATTGTAATTTGTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTGCCATGTGAAAACCTAGAGCATATTG
GAAATGTTCTAAACCTTCTAAAGCTCAGATGCATTGGCATGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTGCACTTAATTGTGGAAAAATATTGCTACAATT
TTAATCTCTGAATGTAATTGCAACTGTGTACATAGCAGGGAGTGTGATCGGGGTGAAATAACTT
GGGCCAGAATATTAAACAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLIKEKICLKLASKLKGVPFGWAMQAAAYIFIHRWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLLQGDFPREIHFHVRYPIDTLPTSKEDLQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSVIPPKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCCACGGCTCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTTGGTTCTCAGTTCTACGAGCTGGTGTAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCCAGGCCTTGGTGGGGAGGACGCCGTGTTCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGGTTCTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGAAGACTGGGAATCTAACGAGATGCCACAGTATCGAGGGAGAAGTGA
GTTTGTAAGGACTCCATTGCAGGGGGCGTGTCTCTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTTCAAGTCCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTCCTCTCATTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTAACTCTGCCTGCTCAGGCTGGTTCCCCAGGCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTGTCTCAGACTCCAGAGCAATGCGAGATGGGAGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAATGCTGGAGCATATGTGTTCCATCCACCTTGTGAGCAGAGTCATGA
GGTGAATCCAAGGTATTGATAGGAGACGTTTCCACCTTGGCCCTGGCTTCTA
TTTACTCGGGTTACTCTGTTGCTGTTGCTGTTGATGGGATGATAATTGTTTCTTC
AAATCCAAGGGAAATCCAGGGGAACACCGCAGTGAGGTGACTCTGGATCAGAGACGGCTCACCCGAAGCTGCG
AGACGCCCGGAAACACGCAGTGAGGTGACTCTGGATCAGAGACGGCTCACCCGAAGCTGCG
TTCTGATCTGAAAATGTAACCCATAGAAAGCTCCCAGGAGGTGCTCACTTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTCTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAATGTAGGGTGGTATGTGGGAGTGTGTTGCTGAGCTAGACAGGGGAAGAAC
ATGTGACTTGTCTCCAAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTGTATTC
ACATTCAATCCCCATTATCAGCCTCCCCCAGCACCCCTCTACAGAGTAGGGTCTTCT
GGACTATGAGGGTGGGACCATCTCTTCTCAATACAATGACCAGTCCCTTATTATACCCCTGC
TGACATGTCAAGTTGAAGGCTTGTGAGACCTATATCAGCATGCGATGTATGACGAGGAAAG
GGGACTCCCATTCTATGTCCAGTGCTCTGGGATGAGACAGAGAACCCCTGCTTAAAGGGC
CCCACACCACAGACCCAGACACAGCAAGGGAGAGTGCTCCCGACAGGTGGCCCCAGCTTCT
CCGGAGGCTGCGCACAGAGAGTCACGCCCCCCTCTCTTCTAGGGAGCTGAGGTCTTCTGCC
TGAGCCCTGCAGCAGCGGAGTCACAGCTCCAGATGAGGGGGATTGGCTGACCTGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGAGCGAATAGACTCACATTAGGTTAGTTGTAAAA
CTCCATCCAGCTAACGATCTGAACAAGTCACAACCTCCCAGGCTCTCATTGCTAGTCACGG
ACAGTGATTCTGCCTCACAGGTGAAGATTAAAGAGACAAACGAATGTGAATCATGCTGCAGGT
TGAGGGACAGTGTGCTAATGATGTGTTTATATTACATTCCCACCATAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTCTATACCAATCACCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTAAAGAATAAAGAGGAGGTAGGATTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACAGGCAAAGAAAACAGAAGAAGAGGAAACTACAGGTCCATATCC
CTCATTAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAATATTTA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGCAGAGTTGGTTAATATTAAAT
ATCAACCAGTGAATTCAAGCACATTAATAAGAAAAAGAAAACATAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCISSLGWFPQPTAKWKGPGQDLSDSRANADGYSLYDVEISI
IVQENAGSILCSIHILAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALGVVVMGMIIVFFK
SKGKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWWYVGCRDDVRGKNNVTLSNNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTGGCAATGATGTATCGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGGTGGGAACCCTGGTTATCGCCTCGTCATCTTCATATCCCTGATTGTCCTGGCAGTGTGCAATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAGAACCTACAATTACTATAGCACATTGTCATTACAAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGCTCTAACATTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTAAAAATGCAATTATAATCTCCATTAAGGGAAATTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAACATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTCACATTGTTACATGAAAGCTGCAAGATGCTGAGGACCCCTAAAGTAG
ATCCTCACTCAGTAAATTAAAAAAATCAACAAAGACAGACAGCTATCTAAACCATTGCTGCCGA
ACACGAAGAAGTAAACTCTAGGTCAAGACTCTCAGGATCGTTGGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGAGTCATCGCTGTTGAGCAACCTTAATTATGCCACATGGC
TTGTGAGTGCTGCTCACTGTTAACACATATAAGAACCCCTGCCAGATGGACTGCTTCTTGGAGTAACA
ATAAAACCTCGAAATGAAACGGGTCTCCGGAGAATAATTGTCCATGAAAAAATACAAACACCCATCACA
TGACTATGATATTCTCTTGCAAGAGCTTCTAGCCCTGTTCCACACAAATGCAGTACATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAGGGTGATGTGATGTTGTGACAGGATTGGAGCACTGAAAAAAT
GATGGTTACAGTCAAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAACTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCTAGAATGTTATGTGCTGGCTCTTAGAAGGAAAACAGATGCA
GCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAAACCCAAACAGGCTGGTGTTTACTAGAGTTACGGCCTGCGGGACTG
GATTACTCAAAACTGGTATCTAAGAGACAAAGCCTCATGGAACAGATAACATTGTTGTTGTTTG
GGTGTGGAGGCCATTAGAGATAACAGAATTGGAGAAGACTGCAAAACAGCTAGATTGACTGATCTCA
ATAAAACTGTTGCTTGATGCATGTATTCTCAGTGTATTCTCCAGCTCTGTCAGTAAGCATCCTGCTCTGCCA
GATCAACTCTGTCATCTGTGAGCAATAGTTGAAACTTGTACATAGAGAAATAGATAACATATTAC
ATTACAGCCTGATTTCATTGTTCTAGAAGTTGTCAGAATTGACTTGTGACATAAATTGTAAT
GCATATATACAATTGAGCAACTCCTTCTCAGTTCTCAGCTCCTCTCATTCAGCAAATATCCATT
TCAAGGTGCAAGAACAAAGGAGTGAAAAGAAAATATAAGAAGAAAAAATCCCCTACATTATTGGCACAGAA
AAAGTATTAGGTGTTTCTTAGTGAATATTAGAAATGATCATATTCATTGAAAGTGAAGTCAAGCAAAGACA
GCAGAATACCAACTCATCATTAGGAAGTATGGAACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
ATATATCCTTATTTCATTCCAAACACTATGATAAATGTGAGAAGATTCTGTTTGTGACCT
ATAAAATTATAACAACTTCATGCAATGTACTTGAAGCAAATTAAAGCAAATATTTTAACATTG
TTACTGAGGATGTCAACATATAACAAATAAAATATAACACCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTNYYSTLSFTTDKLY
AEFGREASNNFTEMSRQLESVKNAFYKSPLREEFVKSQVIKFQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELOSSPVPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:
amino acids 21-40 (type II)

FIGURE 107

AGAGAAAAGAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCGCGAAGAAGTTCCCTG
CCCCGATGAGCCCCCGCGTGCCTCCCGACTATCCCGAGCGGGCGTGGGCACCGGGCCAGC
GCCGACGATCGCTGCCGTTTGCCTTGGAGTAGGATGTGGTGAAGGATGGGCCTCTCCCT
ACGGGGCTCACAAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCCTGCTACGCCCTCAA
TCTGCTTTGGTTAATGTCATCAGTGTGGCAGTTCTGCTGGATGAGGGACTACCTAA
ATAATGTTCACTTTAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTTCT
GTGGTTCATCCGGTCATGATTGCTGTTGCTTCTTATCATTGTTGGGATGTTAGGATATTG
TGGAACGGTGAAGAAGAAATCTGTTGCTTCTGCATGGTACTTGGAGTTGCTGTCAATTCT
GTGTAGAACTGGCTTGTGGCGTTGGACATATGAACAGGAACCTATGGTCCAGTACAATGGTCA
GATATGGTCACTTICAAGGCCAGGATGACAATTATGGATTACCTAGATATCGGTGGCTTACTCA
TGCTTGGAAATTCTTCAAGAGAGTTAAGTGTGAGTATAATTCACTGACTGGTGG
AAATGACAGAGATGGACTGGCCCCCAGATTCTGCTGTGTTAGAAATTCCCAGGATGTTCCAAA
CAGGCCAACAGGAAGATCTCAGTGACCTTATCAAGAGGGTTGGGAAGAAAATGATTCTT
TTTGAGAGGAACCAAAACTGCAGGTGCTGAGGTTCTGGAATTCTCCATTGGGGTGACACAAA
TCCTGGCATGATTCTCACCATTACTCTGCTCTGGCTCTGTTATGATAGAAAGGGAGCCTGG
ACAGACCAATGATGTCCTTGAAGAATGACAACACTCAGCACCTGTCTGACCTCAGTAGA
GTTGAAACAGACTGTCAAGAATCTTGAACACACATCCATGGCAACAGCTTAATACACACT
TTGAGATGGAGGTTAAAAAAGAAATGTCACAGAAGAAAACCACAAACTGTTTATTGGACT
TGTGAATTCTGAGTACATACTATGTGTTTCAAGAAATATGTAGAAATAAAATGTTGCCATAAAA
TAACACCTAACATATACTATTCTATGCTTAAATGAGGATGAAAAGTTCTGATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCTACCCACTGTGAGCC
TGTGTATGACTTTACTGAACACAGTTATGTTTGGCAGCATGGTTGATTAGCATTCCGCA
TCCATGCAAACGAGTCACATATGGTGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTATTA
CTCAGCGATCTATTCTCTGATGCTAAATAATTATATCAGAAAACCTTCATATTGGTGA
ACCTAAATGTGATTTTGTGGTTACTAAATATTCTACCACTAAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTTGTATATAAGTCTGTTAAATCTGTATAATTCACTGCGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAATTGTCCTGTATAGCATCATT
ATTTTAGCTTCTGTTAATAAGCTTACTATTCTGCTCTGGCTTATATTACACATATAAC
TGTATTTAATACCAACTAATTGAAAATTACCACTGTTGAGAATCATTAAACATGTGACAATTAGAGATT
AGAATGATGCTGGCTTGTGAGAAGTATTAAGAAAATTGCACTAACTTAGTTGATTGAGA
AAGGACTGTATGCTGTTTCTCCAAATGAAGACTCTTTGACACTAAACACTTTTAAAAA
GCTTATCTTGCTTCTCCAAACAAGAAGCAATAGTCTCAAGTCAATATAAAATTCTACAGAAAA
TAGTGTCTTTCTCCAGAAAATGCTGTGAGAATCATTAAACATGTGACAATTAGAGATT
CTTGTGTTATTCACTGATTAATACTGTGGCAAATTACACAGATTATAAATTGTTTACAA
GAGTATAGTATATTGAAATGGGAAAGTGCATTACTGTATTGTTGATTTGTTAT
TTCTCAGAATATGAAAAGAAAATTAAATGTCATAAAATTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNTLTAETrVEEAVILTYFPVWHP
VMIAVCCFLIIVGMLGYCGTVKRNLNLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFILGISIGVTQILAMILTITLLWALYYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTACCTCATCCTCATCCTCTTCTCTGATAAAGCCCCTACCAAGTGCT
GATAAAGTCTTCTCGTGAGAGCCTAGAGGCCCTAAAAAAAAAGTGCCTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAAGAGGATTTCCAGTGTCTGGCAGTGGTCCAGAAGGATGCCCTCATTCTGCTTCACCTG
CCTCTCATCACAGGCACCTCCGTGTCACCCGTGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAAGTGGATGAGTCTCAAGGCCTCTATGTGACAACCAGTGAATGGGAG
TGGTACCACTTCACGGCATGGCGGGAGATGCCATGCCTACCTCTGCATACCAAGAAAACACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCAGGGATTGTGCAACGCCAGGTTGTGCCAGCT
TCAATGGAACTGCTGTCTGGAACACCACGGTGAAGTCAAGGCTGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGCTGCTTCCACGCTACTGTGGTCACTTGTGACATCTGCACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAACGTGCTAGGCCCTGACAGGCAGACATGTTG
ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAAGTGAGATCTGTGAAACCTCAAAACTCTACCGCTGT
GAGTGTGGGTTGGCCGTGCTAAGAAGTGTGCAAGACTTGAGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTGCCTGGATCTGAGAAAAGGCTACCAAGTGTGAATGTCCCCGGGCTGGTGTGT
CTGAGGATAACCACACTTGCAAGTCCCTGTGTTGTCAAATCAATGCCATTGAACTGAACATCCCCAGGAG
CTGGTTGGTGGCCTGGAGCTCTCCGTACCAACACCTCTGCCAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTCACTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCGGGGAGCAGCGGGACTTCATCATCGAACAGCAAGCTGCTGATCCCCGTG
ACCTGCGAGTTCCACCGCTGTACACCATTCTGAAGGATACGTTCCAACCTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGATCTCCCAATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTACC
GGGAAGCTCTGCCACCCCAAGCTCGTCACTCCCTACTTGGCATTGAGCCGTGGTGCACGTGAGCGC
TTGGAAAGCTTGGTGGAGAGCTGTTTGCACCCACCTCCAAGATCGACGAGGTCCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAAGATGACTCGTAAGCAGTACACATCCCGGATCACCTAGCAAAGCACTCCAGG
TCCCTGCTTCAAGTTGTGGCAAAGACCACAAGGAAGTGTGACTGCCGGTTCTGTGAGGTG
TTGGACGAGCGTCCCGCTGTGCCACGGGTGCCACGGCGAATCGCTGTGGGAGGAGGAGACTCAGC
CGGTCTACAGGGCCAGCGCTAACAGCGGGCCGATCGCATCGACTGGAGGACTTAGTTCTGAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTTTGGAGCTCTCCCCCACGCCCTCAAGAACATCTGCCAACAGC
TGGGTTCAAGACTCACACTGTGAGTTCAAGACTCCACCAACTCACTCTGATTCTGGTCAATTCAAGTGGCA
CAGGTACAGCACTGCTGAACAATGTGGCTGGGTGGGTTCATCTTCTAGGGTTGAAAACAAACTAACTGTCCA
CCCAGAAAGACACTACCCCAATTCCCTCATTTCTTCTACACTTAAATACCTCGTGTATGGTGCACAGAC
CACAAAATCAGAAGCTGGTATAATATTCAAGTACAAACCCTAGAAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTAAATATGTAATTAGTTACCTTGAAATTCAATTCAAGTGCAGACTAA
TTATAGGGATTGGAAGTGTATCAATAACAGTATATAATT

FIGURE 110

MPPFLLLTCFITGTSVSPVALDPCSAISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYYVYRLTKPSVCFHV
YCGHFYDICDEDCHGCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSRGKTCEDVEGCHNNNGGCSHSLGSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFFLTNTSCRGVSNGLHVNILFSLKTCGTVVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEEPYREALPTLKLRSLSYFGIEPVHV
SGLESLVESCFAPTSKEVLKYLIIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTGCTCAGGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGTGACCAGGGACTTCTGACCTGCTGCCAGCCAGGACCTGTGTGGGAGGCCCT
CCTGCTGCCCTGGGTGACAATCTCAGCTCAGGCTACAGGGAGACCGGGAGGATCACAGGCCAGCATGT
TACAGGATCCTGACAGTGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGAAACCCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGCCTGGGAGGTATCATCATTGT
GGTTGTCCCTCATCAAGGTGATTCTGGATAAAACTACTTCCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCCGCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAACTGGTTCTCTGCCTGTTGACAACCTCACAGAAGCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGGAGATTGCCAGACCAAGGATCTGGATGTTGAAATCACAGAAAAGCCAG
GAGCTTCGCATGCCAAGTGGCCCTGTCTCTCAGGCCCTGGTCTCCCTGCACTGTCTGCCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGGAGGAGGCCCTGTGGATTCTGGCCTPGCAGG
TCAGCATCCAGTACGACAAACAGCACGCTGTGGAGGGAGCATCCTGGACCCCCACTGGCCTCACGGCA
GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGCAG
CTTCCCCTCCCTGGCTGTGGCCAAGATCATCATATTGAATTCAACCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTCTCAGGCACAGTCAGGCCATCTGTCTGCCCTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATTGGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGTCAGACATACTGCTGCAGGGTCAGTCCAGGTATTGACAGCACACGGTCAATGCAGACGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGATCGTTAGCTGGGCTATGG
CTGCGGGGCCAGCACCCAGGAGTATACACCAAGGCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCTTGCACTGCTGGAGGCCCTCCCTGCCCTGCCACCT
GGGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTGGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGCCCTCAATTCTGTAAGAGACCCCTCGCAGCCCAGAGGCCAGAGGAAGTCA
GCAGCCCTAGCTCGGCCACACTTGGTGTCCCAAGCATCCAGGGAGAGACAGCCCACGTAAAGGTCT
CAGGGTATTGCTAAGCCAAGAAGGAACCTTCCACACTACTGAATGGAAGCAGGCTGTCTGTAAAGCC
CAGATCACTGTGGCTGGAGAGGAAGGAAGGGCTGCCAGCCCTGCGCTTCAACCCATCCCCAA
GCCTACTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCCTACCT
ACTGTTGTCATTGTTATTACAGCTATGCCACTATTATAAGAGCTGTGTAACATCTCTGGCAAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVVLIKVILDKYYFLCG
QPLHFI PRKQLCDGELDCPLGEDEEHCVKS FPEGPAVAVRLSKDRSTLQVLD SATGNWFSACFDN
FTEALAE TACRQMGYSRAVEIGPDQDLDVVEITENSQELRMRNSSGPCLSGSLVSLHCLACGKSL
KTPR VVGEEASVDSWPWQVS IQYDKQHVCGGSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKL
GSFPLSLAVAKIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICL PFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCQGDGGPLMYQS
DQWHVVGVIVSWGYGCCGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCAAGTGATCCACCCGCCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGCCCTCTGAACAACCTTTCACTGCAACTAAAAAGCCACAGGAGT
TGAACGTGCTAGGATTCTGACTATGCTGTGGCTAGTGCTCCTACTCCTACCTACATTAAATC
TGTTTTTGTCTCTGTAACTAGCCTTACCTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACTCTGGAACAGAACAGAGGTTCTACCCACACCGTCCCCTGAAG
CCGGGGACAGCCTCACCTGCTGCCCTCGCTGGAGCAGTGCCCTACCAACTGTCTACGTCT
GGAGGCAGTCACTCGGGCAGTGAGGTAGCTGAGCCTCTGGTAGCTGCGGCTTCAAGGTGGC
CTTGCCTGGCGTAGAAGGGATTGACAAGCCCAGAAGATTTCATAGGCGATGGCTCCACTGCC
AGGCATCAGCCTTGCTGTAGTCATCACTGCCCTGGGCCAGGACGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACCGCTGCCGCCATCTAACCTTCTATGTCCTGCACATCACCTG
ATCCATGGGCTAATCTGAACCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTGGAAACATGAGG
GAACGCCGAGGAAAGCAAAGTGGCAGGGAAAGGAACCTTGTCAGAAATTATGGGTAGAAAAGATG
GAGGTGTTGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGAGCTGACATGTGGGGAAAGGCTG
CCGATGGCGCATGACACACTCGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCCGAT
CCACGTACAGCTGCTGAAGGGCAACTGCAGGCCATGCTCATCAGCCAGGCAGCAGCCAAA
TCTGCGATCACCAGCCAGGGCAGCCGTGGAGGAGCAAGCAAAGTGACCTTCTCC
CTCCTTCCCTCTGAGAGGCCCTCTATGTCCTACTAAAGCCACAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTAGCAAGGCCAGGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGGAAATGCCCTCAGTAAGCACAGGCTGCAAATCCCCAGGCAAAGGACTGTG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAGGAGCTAGAGCTTGGTT
CAAATGATCTCCAAGGCCCTTATACCCAGGAGACTTTGATTTGAAATTGAAACCCCAAATCCA
AACCTAAGAACCGAGGTGCAATTAGAACATCAGTTATTGCCGGTGTGGCTGTAATGCCAACAT
TTGGGAGGCCAGGCCGGTAGATCACCTGAGGTAGGAGCTCAAGACCAAGGCCCTGGCAACATGG
TGAAACCCCTGTCTACTAAAAATACAAAAAAACTAGCCAGGCATGGTGGTGTGCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAAATTACTTGAAACCTGGAGGTGAAGGGAGGCTGAGACA
GGAGAACACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTA
TGGTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKVFCSLVTSLYLPNTEDSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGCCTCAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAAATTGTCCT
GTTTGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGAAAGTGCACGACTTAAAAACGGATAACACTGG
CATCTACTTCGTGGTCTCAAAAATGTTTATCAAAACTCAGATTAAGTGATTCTGAATT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTCTTGAAACAGTCAGTG
ATTGGGTCCCAGCAGAAAGCCTATTGAAAACCGAGATTCTTAAATTCCAAAATTCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACCTAATATCAGTTCTGAGTTACAAG
ACTTTGAGGAGGGAGGAGAAGATCTTCACTTCCCTGCCAACGAAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTCCAATAATGACTATACTGAAAATGGAATAGAATTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCTGTGAACCTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGCGTGTATCATGCCCTG
TAACCTGGTGGTGGCCCGATGCTGGGGAGGGCTTAATAGGAGGTTGAGCTAAATGCTTAAAC
TGCTGGCAACATATAATAATGATGCTATTCAATGAATTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCAGAATTACTGTAGGTAATTCTCTCTCATGTTCTAATAAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKKICKSLKICGLVFGILALTILVLFWGSKHFPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVE
LQDFEEEEDLHFPAKEKKGIEQNEQWVVPQVKVEKTRHARQASEEEELPINDYTENGIEFDPLD
ERGYCCLYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTTCACACCTCGCAGCAGGAGGGCGGCAGCTCTCGCAGGGCA
GGCGGGCGGCCAGGATCATGTCCACCACACATGCCAAGTGGTGGCGTCTCTGTCCATCCTGGGCT
GGCCGGCTGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTGGAGGAGCTGCGTGAAGCAGAGTTCAAGGCTCACCGAATGCAGGCC
TATTTCAACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCC
GGGTGCCATTGGCCTCTGGTATCCATCTTGCCTGAAATGCATCCGATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGATCATGTTCAATTGTCTCAGGTCTTGCAATTGCTGGAGTG
TCTGTGTTGCAACATGCTGGTGAACATCTGGATGTCACAGCTAACATGTACACCGGCATGGTGG
GATGGTGCAGACTGTTCAAGACCAGGTACACATTGGTGCAGCTGTTGTGGCTGGTCGCTGGAGGCC
TCACACTAATTGGGGTGTGATGATGTCATGCCCTGCCGGCTGGCACCAAGAAACCAACTACAAA
GCCGTTCTTATCATGCCCTCAGGCCACAGTGTGCTACAAGCCTGGAGGCTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGACTATGTGTAATGCTCTAACACCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCACATCTAGATTCTCTTGACTCACAGCTGGAAAGTTAGAAAAGCCT
CGATTTCATCTTGGAGGGCAAATGGCTTAGCCTCAGTCTGTCTCAAATATTCCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGTATAGCTCACATTCAATCCTCTATTCTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTAATCTCTCTCACATTGATGATTTAGACAGACTCCCCCTC
TTCCCTCTAGTCATAAAACCCATTGATGATCTATTCCAGCTATCCCCAGAAAACCTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTTCTGCTGTTGAATTGCTCTCCCCACCCCAACTTGGCTAGTAATAA
ACACTTACTGAGAAGAAGCAATAAGAGAAAAGATATTGTAATCTCTCCAGCCCATGATCTGGTTTCTT
ACACTGTGATCTAAAAGTTACCAACCAAGTCATTTCAGTTGAGGCAACCAACCTTCTACTGCTG
TTGACATCTCTTATTACAGCAACACCATTCTAGGAGTTCTGAGCTCTCCACTGGAGTCCTTTCTGT
CGCGGGTCAGAAATTGCTCTAGATGAATGAGAAAATTATTTTTAATTAAAGTCCTAAATAGTAA
AATAAATAATGTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTGACATTGCTATATGGTACTTGTAAAGTCATGCTTAAGTACAAATTCC
ATGAAAAGCTCACACCTGTAATCCTAGCATTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCACAAAATACAGAGAGAAAATCAGCCAGTC
TGGTGGCATAACCTGTAGTCCCAGCATTCCGGAGGCTGAGGTGGAGGATCACTTGAGCCCAGGGAGGT
TGGGGCTGCACTGAGCCATGATCACACCACTGCACCCAGGTCAGATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAAACTAATTCTTAA

FIGURE 118

MSTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGVWAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFSNTKNKKIYDGGARTEDEVQSYP SKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTCTTGTGGCACAGAGAACCTGCTCAAAGCAGAAGTAGCAGTCCGGAGTCC
AGCTGGCTAAACTCATCCAGAGGATAATGGCAACCCATGCCTAGAAATCGCTGGCTTCTTGTGGTGTGGAATGGCACAGTGGCTGTCAGTGTCA
GCTCAGTGGAGAGTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCACAAATCTATGATTCCCTGCTGGCTTTCTCCGGACCTACAGGCAGGCCAGAG
GACTGATGTGTGCTGCTCCGTGATGTCCTCTGGCTTCTATGATGGCATCCTGGCATGAAATGC
ACCAGGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGCTCATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGTGATTGTTGGAGGAGCTCTGTTCTGCGTTTTGTTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTTCCCATCGCACAACCCAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTTAGTTGTGTATGTTTTAACTTACTATAAAGCCATGCAAATG
ACAAAAAATCTATATTACTTTCTCAAAATGGACCCAAAGAAACTTGATTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAACGTGCATCAGCTATTATGATTCTATAAGCTATTCTCAGCAGAATGAGATA
TTAAACCCAATGCTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGGTCAAGCATTCTA
CTCTTTTATCATTACTTCAAAATGACATTGCTAAAGACTGCATTATTTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTTATATGGT
TTTATTAAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTGTTAAAACAGCTTAGGGATTATGCTTCCA
TTTATAATGAAGATTAATGAAAGGTTAATCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCCTAATTCTTATCCTCTCCAGAGGCTTTTT
CTTGTGTTAAATTAAACATTGTTAAAACGCGAGATATTGTCAGGGCTTGCATTCAAACGTCT
TTCCAGGGCTACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGTTAGGAAAGTG
AAAATTTTGTGTTGTGTTGAAAGAAGAATGATGCATTGACAAGAAATCATATATGTATGGAT
ATATTTAATAAGTATTGAGTACAGACTTGAGGTTCTCAATATAAATAAAAGAGCAGAAAATA
TGTCTGGTTTCATTGCTTACCAAAAAACAAACAACAAAAAGTTGTCCTTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCACATTGTCATTGTTCTGTGAAAATAAATTCTTCTGTA
CCATTCTGTTAGTTACTAAATCTGAAATACTGTATTCTGTTATTCAAATTTGATGAA
ACTGACAATCCAATTGAAAGTTGTGTCACGTCTGTAGCTAAATGAATGTGTTATTGCTT
TATACATTATATAAATAAATTGTACATTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSIFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWANAIIRDYNSIVNVAQKRELGEALYLGWTALVLIVGGALFCCVFCCNEKSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGGCGCGGGGTGAAAGGCGCATTGATGCAGCCTCGGCGGCCTCGGAGCGCGGCGAG
CCAGACGCTGACCACGTTCTCTCGGTCTCCTCCGCTCCAGCTCCCGCTGCCAGCC
GGGAGCCATGCGACCCAGGGCCCCGCGCCCTCCCGCAGCGGCTCCGCGGCCCTGCTGCTCC
TGCTGCTGAGCTGCCGCGCCGAGCGCCTCTGAGATCCCAAGGGGAAGCAAAGGCGCAG
CTCCGGAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGCCAATGTTATTCCGGGTACACCTGGATCCCAGGTGGGATG
GATTCAAAGGAGAAAAGGGGAATGCTGAGGGAAAGCTTGAGGAGTCCTGGACACCCAACATAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTTGGAAAATTGCGGAGTGTCATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTGTTCACTGGCTCACTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTACATTCAATGGAGCTGAATGTTCAGGACCTCTTCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAAATTATCG
CACTTCTCTGTGGAAGGACTTGTGAAGGAATTGGTGGTGGATTAGTGGATGTTGCTATCTGG
TTGGCAGTTGTCAGATTACCCAAAAGGAGATGCTTCACTGGATGGAATTCACTGGTGGCTCGC
ATTATTGAAGAACTACCAAAATAAATGCTTAATTTCATTGCTACCTCTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTAAATAAGTTATGTATACTGAAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAAAGTGTGATTCACACTGTTAAATCTGATTATTCACTTGG
CTTCAATCAAAAGGGTTCAATATTTTTAGTTGGTTAGAATACCTTCTCATAGTCACATT
CTCTCAACCTATAATTGGAATTGTTGGTCTTTGTTCTTAGTATAGCATTGTTA
AAAAAATATAAAAGCTACCAATTGTTACAATTGTAATGTTAAGAATTTTTATATCTGT
TAAATAAAATTATTCCAACA

FIGURE 122

MRPQGPAA SPQR LRG LLLL LQLP A PSS A S E I P K G K Q K A Q L R Q R E V V D L Y N G M C L Q G P A G V P G R
D G S P G A N V I P G T P G I P G R D G F K G E K G E C L R E S F E E S W T P N Y K Q C S W S S L N Y G I D L G K I A E C T F T K
M R S N S A L R V L F S G S I R L K C R N A C C Q R W Y F T F N G A E C S G P L P I E A I I Y L D Q G S P E M N S T I N I H R T S
S V E G L C E G I G A G L V D V A I W V G T C S D Y P K G D A S T G W N S V S R I I I E E L P K

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGCTCTCTGCCTCTGGCTCCAACGCAGCTGTGGCTGAA
CTGGGTGCTCATCACGGGAACGTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATACATCATGTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCCAAAAAACTGTAAGAGATGCAAAAAGCTAATATCCATGAAGATCC
TATTACCTAGGAAGATTGTGTTGTCGAATGCGGTGTTGGGATTATTGTTCTTGGAG
TGTCTGCGTGGCTGGAAAGAATAATGTTCAAATCGGTCCATCTCCAAGGGTCCAATTT
TCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGCTGTCATGCAACTG
GCCCTAAGCCAAAGCAAAAGCCTAAGGACGACCTTGAACAATACAAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGAGCACTGGTTAGCCCCACTGTCCTACTGACAATG
CTTCTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGAAGGCAAATGTTATTGTGA
ATCTCAGAAATTACAGGAGATACTCCTCAAGTATACTGCTGGTTGCTTAGGTTGTCCTTCGCT
ATAACAGCCTTCAAAACTTAAGTATACTCAATTAAAGGGCTCAACAGCTCACCTGGCTATAC
CTTGACCCATAACCATACTCAGCAATATTGACCAAAATGCTTTAATGGAATACGCAACTCAAAGA
GCTGATTCTTAGTCCAATAGAATCTCTTAAACAAATCCTCAGACCTGTGACAATT
TACGGAACCTGGATCTGCTATAATCAGCTGCATTCTGGATCTGAACAGTTGGGGCTTG
CGGAAGCTGCTGAGTTACATTACGGCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
AGACTGGCAACCTGGAACCTGGACCTGGATATAACCGGATCCGAAGTTAGCCAGGAATG
TCTTGCTGGCATGATCAGACTCAAAGAACCTCACCTGGAGCACAATCAATTTCAGCAGCTAAC
CTGGCCCTTTCAAGGTTGGCAGCCTCAGAACCTTACTTGCACTGGAATAAAATCAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGTCCGAATCTGAGCGCTCAACCTGGAT
TCCAACAGCTCACATTATTGGTCAAGAGATTGGATTCTGGATATCCCTCAATGACATCAG
TCTTGCTGGAAATATGGGAATGCAAGCAGAAATTTGCTCCCTGTAAACTGGCTGAAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTCAGTCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCACTGAAGAAACTACAGCATCTGGACAAAGTACTACAGAGAGGTTGATCTGGCAG
GGCTCTCCAAAGCCGACGTTAACGCCAACGCTCCCCAGGCCGAAGCATGAGGAGAACCCCCCT
TGCCCCCGACGGTGGGAGCCACAGAGCCGCCAGAGACCGATGCTGACGCCGAGCACATCT
TTCCATAAAATCATCGCGGGCAGCGTGGCTTTCTGTCCGTGTCATCCTGCTGGTAT
CTACGTGTCATGGAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAAGCGCCTCATGCGAA
GGCACAGGAAAAAGAAAAGACAGTCCCTAACGAAATGACTCCAGCACCCAGGAATTATGTA
GATTATAACCCACCAACCGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCTGCACCTA
TAACAAATGGGCTCCAGGGAGTGTGAGGTATGAACCTGTGATAAAAAGAGCTTAAAGCT
GGGAATAAGTGGTGTCTTATTGAACTCTGGTACTATCAAGGAAACGCGATGCCCTCC
TTCCCTCTCCCTCACTTTGGCAAGATCCCTTGTCCGTTAGTGCATTATAACT
GGTCATTTCCTCTACATAATCAACCCATTGAAATTAAATACCAATCAATGTGAAGCTT
GAACTCCGGTTAATATAATACCTATTGTATAAGACCCCTTACTGATTCCATTATGTCGATT
GTTTAAGATAAAACTTCTTCATAGTAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNNTFR
PVTNLRNLLDLSYNQLHSLGSEQFRGLRKLLSLHRSNSLRTIPVRIFQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLHEHNQFSKLNALFPRLVSLQNLQWNKISVIGQTMSTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNSNLTFIGQEILDWSISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKKRQSLKQMTPSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATGTCCGTCCGGAGGAGGGAGGCTTTGCCGCTG
ACCCAGAGATGGCCCGAGCGAGCAAATTCTACTGTCCGGCTGCGCGCTACCGTGGCGAGCT
AGCAACCTTCCCCCTGGATCTACAAAAACTGACTCCAAATGCAAGGAGAAGCAGCTTGTCTC
GGTTGGGAGACGGTGCAAGAGAACTGCCCCCTATAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTAACAGACACGTAGT
GTATTCTGGAGGTCCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
TTAGCCAATCCAACGTACCTAGTGAAGGTTAGATGCAAATGGAAGGAAAAGGAAACTGGAAGG
AAAACCATTCGATTCTGTGGTGTACATCATGCATTGCAAAATCTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGGGCTGGTACCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTAA
ACCACTTATGATACTGAAACACTACTTGTATTGAATACACCACCTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTATGTTCTGGACTGGTAGCTTCTATTCTGGAACACCAGCGATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAGGAAGGGACTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGTGTTCAAGGTGAAGGATTGAGTCATGAGTCATATAAAGGTTTAC
ATCTTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACTTATGAAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLCVVFGKSEDEHYPLWKS
VIIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCAAGCAGGTGGCGGCGGCAGGAGAGCGGCCGGCGTCAGCTCCTCGAC
CCCCGTGTCGGCTAGTCCAGCGAGGCGACGGCGGCGTGGGCCATGGCCAGGCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCCTGGTACGGGGCCTCGGGGGCATCGGCGGGCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGCTGCGCCCGACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTCAGGCTACCCCGGACTTGTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTAGCTATCCGTTCTCAGCACAGCAGGTAGACATC
TGCATCAACAATGCTGGCTTGGCCGGCTGACACCCCTGCTCAGGCAGCACCGAGTGGTTGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGAAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGCACATCAATAGCATGCTGGCCACCGAGTGT
CCCCGTCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCCAGGTGTTGG
AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCTGAGAAGGCAGTGCACCTATGAGCAA
ATGAAGTGTCTCAAACCGAGGATGTGGCCAGGCTGTTATCTACGTCTCAGCACCCCGACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGAGCTCC
TCCTCCCTCCCCACCCCTCATGGCTGCCCTGCCTGGATTTAGGTGTTGATTCAGTGGAT
CACGGGATACCACTCCGTCCACACCCGACCAGGGCTAGAAAATTGTTGAGATTTATA
TCATCTGTCAAATTGCTTCAGTTGAAATGTGAAAAATGGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTTGCCCTGGCACTGGCCTTGTCTGCTCTCAGTG
TCTTCCCTTGACATGGGAAAGGAGTTGTGGCCAAATCCCATCTTGCACCTCAACGCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCTCACCTATATCTGTGTTATCCAGGGCTCC
AGACTTCCCTCCTGCCTGCCCACTGCACCTCTCCCCCTATCTCCTCTCGGCTCCCC
AGCCCAGTCTGGCTTCTGTCCCCTCTGGGTCACTCCCTCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCTGGCCAGTGGATTTCATGGTGATCATTAAAAAGAAAATCGAACCAA
AAAAAAA

FIGURE 128

MARPGMERWRDRRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDLNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRLPLSVTHFYSATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGGTGTCTTCCTCAGCCTCCTGCCGGTGCCTACACCAT
CATGTCCCTCCCACCCCTCTTGACTGCGGCCGTTCAAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCTCCGAGGCAGTCTGCTCAGAGGCCTCGGCCAGAATCCAGTTCTGGTTCATGC
CAGCCTGAAAAGGCCATGGAACTTGGGTGAATCACCAGTGCCTTAAGAGGGTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTCTGCCTGTTCAAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAAATGAGGAACGTGAGAAAATTAAATTCTCATGTATTTCATTTA
TTAATTAAACTGATAGTTGACATATTGGGGTACATGTGATATTGGATACATGTATA
TACAA
TATATAATGATCAAATCAGGGTAACCTGGGATATCCATCACATCAAACATTATTTTATTCTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTCAAGCGATTCTCATGCCCTCACCTCCAAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCAACTAATTTTGTATTAGTAGAGACGGGTTTGCCATGTTGCCAGG
CTGGCCTGAACTCCTGGCCTCAAACAATCCACTTGCCCTGGCCTCCAAAGTGTATGATTACA
GGCGTGAGCCACCGTGCTGGCTAAACATTATCTTGTGTTGGAACTTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTCCCTCT
ATCTAACTGTATATTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTATCCTCCC
AACCTCTGATCACCTCATTCTACTCTACCTCCATGAGATCCACTTTTAGCTCCACATGTG
AGTAAGAAAATGCAATATTGTCTTCTGTGCTGGCTATTCACTAACATAATGACTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTCGTTCTTAATTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSPRGSSLRGPRPRIPVLVSCQPV
KGHGTLGESPMPPFKRVFCQDGTVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTCGCTGCGATGGGGTTCATTTACTGTAATGATCGCTTCTGACATCCATTCAAACAG
GAATACCAGAGGATGCTACAACCTCTCACCTCAGAACACAATAAATAATGCTGGATTCCCTCAGAT
TTGAAAAAAACTTGTGAAAGTAGAAAGAATATACTATACCAACAGTTAGATGAATTCTACCAACCT
CCCAAAGTATGTAAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCCTATCTGAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCGACTGCTTTCTGTCCCGTAATCACCTAGCACAAATTCCCTGGGGTT
GCCAGGACTATAGAAGAACTACGCTGGATGATAATCGCATATCCACTATTCATCACCCTCTCAAG
GTCTCACTAGTCTAAACGCCTGGTCTAGATGAAACCTGTGAACAATCATGGTTAGGTGACAAGATT
TTCTTCAACCTAGTTAATTGACAGAGCTGTCCTGGTGCAGGAAATTCCCTGACTGCTGCACCAGTAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATAACCACATCAATCGGGTGCCTCAAATGCTTTT
CTTATCTAAGGCAGCTATCGACTGGATATGCTCAATAAACCTAAGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAACGATTCTCGCAACAATCCCTGGTATTGCGGGTGCAGATGAAATG
GGTACGTGACTGGTACAATCACTACCTGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCAACTGCAATACCCAAACACAGTGTATCCTGCCAAGGACAGTGGCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCAAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTCTATCTTGAAACTTGCTCTACCTATG
ACTGCTTGAGACTCAGCTGGCTTAAACTGGCCATAGCCCGCATGGATCTATAACAGAAACAATTG
AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTCACCTATAAAGTATGCATGGTC
CCATGGAAACCCAGCAACCTCTACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCT
CGAATGTACAACCTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTAC
TTGGCTGCCATCATTGGGGCTGTGGCCCTGGTACCTGCTTCTTGCTTGTGTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTCTCTCAAGGAACCTGCAAGGAGACTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCCTGGAAATCAGGGAAACTCTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTGAAGGAGGAGTTGTAATACACACCATATTCTCTCTAATGGAATGAATCTG
ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTGGTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDVKFFNLVNLTESLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQYRLDMSNNNSNLPQGIFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNRGILMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQTIAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQQTTGSPSRKTITITVKSVDIISWKLALPMTALRLSWLKGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVMVPMETSNLYLFDETPVCIETETAPLRMYNPTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKD
NSILEIRETSFQMLPISNEPISKEEVIHTIFPPNGMNLYKNNHSESSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCCAGAGTCCTTGCAGGCCACCCAGGCTTCTGGCA
GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGCTTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGCTGCTGCTCTGCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGGCCACCGTCTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGCCTGGAAGAGCAGCTAAAAGTGTCCGGAGAAGTGGGCCCAGGAGCCCCCTGC
TGCAGCCGCTGAGCCTGCGCTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGTACCTAACAGAGAGTTCCAGCTGACCCCTCCAGCCGG
GTTTGAAGCTCCACATGCCTGGATCCACACTGATGCCCTTGGTGTACCCACGTTGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGTGACGTGCTGGTGAGCTGCTGGGAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCTGTCACCAACTGCTCTTCTCTGGGCCAGAATGAGGGATGCACACAGG
GACCACCTCAACAGAGCCAGGACTATATCAACCTTCTGCGCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCAACCTACCCGACATCTTCTGGAAAACATCATGTTCTG
TGGAATGGCGGCTCTCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTCGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGACTGAAGAGGCAGAAAAACATTCCAGATTCTCGCTCTG
TGCTCAGGCTGGAGTACAGTGGCGCAATCTGGCTCACTGCAACCTTGCCTCTGGTTCAAGC
AATTCTTGCCTCATCCTCCGAGTAGCTGGACTACAGGAGCGTGCCACCACGGCTAAT
TTTATTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCGGCCACCTCAGGCTCCAAAGTGTGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAAGGGCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCTCCCCTAGATGGCTGCTCTCCCACAAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCCCTACATCCTGGCAGAATACCCCCCAGCAAACAGAGAGGCCACACCCATCCACACCG
CCACCAAGCAGCCGCTGAGACGGACGGTCCATGCCAGCTGCCCTGGAGGAGGAACAGACCC
TTAGTCCTCATCCCTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTCAACTGCAACTGAAAAAAAAAAAA

FIGURE 134

MSARGRWEGGGRRACRGSLGLARAQGAERVTSSEQRPMASLGLLLLLTLPLWSSSLPGLD
TAESKATIADLILSALERATVFLERQLPEINLDGMGVVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQTLQPGFWKLPWAHIDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSSPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGLSLQPLPPGFQFSCLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTCCCGCTGCTGCTG
CTGCTGCTATGGGGATGTCCAGAGCTGGAGGTGCCCGGGCTGCTGCTGAGGGATCGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTCAAGATTGAGGGCGTGCAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTGGCGCCGAGTGTGGTAGACGGAGAAGAGCAGTCGGTTCTTAAG
ACAGATGGGAGTTTGTGGTTATGATATACTTCTGGATCTTATGTAGTGGAAAGTTGATCTCC
AGCTTACAGATTTGATCCGTTGAGTGGATATCACTTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCAAATGAAATCTCAGGT
CCACCTTCTTACTTTATTAAAAGGGATCGGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTCTTCCTTATTGATATTGTGCTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTCTCTCAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
AAAAACAGGCAAAAGTGGGCTGGAAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGCAC
AAACACGGCAACACTGGTGGCATCCAAGTCTGGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCGACGTTGATCTTACAACGTGTATGTT
AACTTTTAGCACATGTTGTACTTGTACACGAGAAAACCCAGCTTCATCTTGCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAATTAT
ATGAACTACTACATTATGTATATTAATTAAAACATCTTAATCCAGAAATCAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSSEVPGAAEGSGGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGFLLKTDGSFVVHDIPSGSYVVEVVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMMVPLLIIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTCGATTCTCCCCAGTTCCCTGTGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAACAA
GTTTGACATTCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGACTAACGGCGAGCCAGGGATGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTCTTCACGGGAGGCTTGGCAGT
TTTTCTTACTCCTGTTGCTCCAGATTTCAGGCCTAAGATGAAAGCCTTAGTCTTGCCTCAGC
CTTCTCTGCTGCGTTTATTCCTATGGACTCCTTCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTCAGGAAATACGAAATGGATTCTGAGATAACGGGCAGTG
TGCAAGCCAAAGATGAAACATTGACATCAGAATCTAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCTCGGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTACTCTCCGAAGATCAGCAGCCTGCCAATTCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAAATACGCCAGATTCTGAGTCATTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTTCAATACCTGAGAGGAGGATGACCCCAA
ACCACCATCTTTACTGTACTAGTCTTGCTGGTCACAGTGTATCTTATTGCTTACTTG
CTTCCTGCTGATGATTGTCTTATGCATCCCCAATCTAATTGAGACCATCTGTATAAGATTT
TGTAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTATTGCTATT
ATGTATTATTTTACTTGGACATGAAACTTAAAAAAATTCACAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTATACAGTAAAAAAACCTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCAATTGCTTCAACTAAGGACATATTACTCATGCTGATGCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTTGAAATAAGTTGATGTTGAAATTGCAC
ATCTACCTTACAATTACTGACCACCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTCTTGCTACCAAAAAAA
AAAAAA

FIGURE 138

MRQFPKTSFDISPMSFSIYSLQVPAVPGLTCWALTAEPGWQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIKKDLRLSHAHMTCHCCEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGAAGCGCGCTGCAGCAGGGGAGGCTCAGGTGGGTCGGTCCGCATCCAGCC
TAGCGTCCACGATGCGGCTGGCTCGGGACTTCGTACCTGTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTCCTCGGGATTCTCCCGCTCCGCTGGCTGGAGCCAGTTCAACTGGACCACGCTGCC
CACGGAGCGGAGCCCCAGGCCGAACCTCGGCTGGAGCCAGTTCAACTGGACCACGCTGCC
ACCACCTCTTCAGTAAAGTGTATTGTTCTGATAGATGCCCTGAGAGATGATTTGTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTTACGGAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAGCAGGACCTACAGTTACTATGCCCTGAATCAAGGCATTGATGACGGGGAG
CCTTCCTGGCTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAGAGACAGTGTGA
TAAGACAAGCAGCTGGAAAAGAATAGTCTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAAGCATTGTGGAATATGATGGAACAACCTCATTTCGTTGTCAGATTACACAGAGGT
GGATAATAATGTACAGGAGATTGGATAAAGTATTAAGGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGCTGGACCACATTGGCCACATTTCAGGGCCCAACAGCCCCCTGATTGGCAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGA
GACGCCCTTACCCAATTGCTGTTCTTGTTGACCATGGCATGTCGAAACAGGAAGTCACG
GGGCCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTAATCAGTTCTGCGTTGAAAGGAAA
CCGGTGTATATCCGACATCCAAGCAGTCCAATAGAGGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCTATTCCAGTTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCACATTGAATACAGTGCAGCTAGTAAACTGTTGCAAGAG
AATGTGCCGTATATGAAAAGATCCTGGGTTGAGCAGTTAAATGTCAGAAAGATTGCATGG
GAACGTGATCAGACTGTACTTGGAGGAAAAGCATTCAAGGTCTATTCAACCTGGGCTCCAAGG
TTCTCAGGGCAGTACCTGGATGCTCTGAAAGACGCTGAGCTTGTCCCTGAGTGCACAAAGTGGCCAG
TTCTCACCCCTGCTCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCA
CTGTCATCTCCGGGTTCTCTGCTCTTATTGGTATCCTGGTTCTTCGGCCGTTACGT
CATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTCGCTCTGTCGGCTGGCGAGGCT
GCCTTCGTTACAGACTCTGGTTGAAACACCTGGTGTGCCAAGTGCCTGGCAGTGCACACTGG
AGGGGGCCTCAGGAAGGACGTGGAGCAGCCTATCCCAGGCCTCTGGGTGCCCCACAGGTG
TTCAACATCTGTGCTCAGGTCAAGATGCCCTAGTTCTGGAAAGCTAGGTTCTGCGACTGTAC
CAAGGTATTGTAAGAGACTGGCGTCACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAA
TCGGACAGCCTCCAGCAGAGGTGTGGAGCTGAGCTGAGGGAGAGACAATGGCCTGG
CACTCAGGAGGGTCAAAAGGAGACTTGGTGCACACTCATCCTGCCACCCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTCTTCCAAGGGGACGTTCTGTTGAAATTCTAGTCCTGGCC
TCGGACACCTTCATTGTTAGCTGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGGCGATGGTCAC
ACTCAGATCCACAGAGGCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCC
ACCCCAACCCCTGCACAGCCCTCATCCCTCTGGCTTGAGCCGTCAGAGGCCCTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTCATCAGGGCACAGGCTTCTCGGAGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCCATCTGGGCTATGCTCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFVSDYTEVDNNV
TRHLDKVLKRGDWDLILYLHGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAACT
TCCCTATAGAAAACACTGCCACCTTAAGACCACTACACCTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTCATGACCAGGATCACAAAGTACTGGTCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAAACTACATACGCCAGAGATCTTCTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGAGTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAAACTGATGAAGCT
GGCTGCCAAAAGGAATCAGCACGCCGCCCTCATCTTATAGGGCTCAGGTGGCTCCTGGA
ACATGCTGGAGTCGGCGCTCACCCGGATGGTCATCTGCACCTCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTGAGAACAGGAAACACATTGAATTTCATTCAACCAGTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCCTGAACGCCCTCGCTA
ATTGAACTATTGTATAAAAACACCAAACTGCTCACT

FIGURE 142

MLLLLLEYNFPIENNQHLKTTHTFRVKNLNPKKFSIHQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCCTCCCTTAA
TCCAGGATCCTGTCCTCCTGTCCTGTTAGGAGTGCCTGTCAGTGTGGGTGAGACAAGTTG
TCCCACAGGGCTGTCAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGCCTGCCGTGG
AAGGGAGGTCTGTCCTGGCGCTGCTGCTGCTCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTGCTGTCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGAAATTGAAGACGACATTGACAAC TGCCATTCCAAGAAAGCACAG
AGCTGAACAATACTTACCTGCTTCTTCACCACAGCACCCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTGAGTGAACACCACACTCACAGGCTGTCCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTGTTAGAAAATCC
ACACATCTTGAACCTAATCATGTAGTGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGCTGGCACTGTCTGCATGCTGCTCTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCAGCAGCTGCCAGATCA
CTGAGGGCCCAGGTGGCTGAGAACCGCCGGAGCCTCATCAAGCAAGGCCAGCTGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACAGAGGCCAACTACTGGCAGTTCCCAGTGGCATCCA
CTACAAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTACCGGCTGCATCAATGCCA
CCCAGGCGCGAACCAAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTGGTGGAGAGGGCGCAGGACT
TCGGGTACCATGCAACCAGCCAGTGCTCCTGCCTCTGGCTTGATCTGGCTATGGTGAAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGGCCAGCAGCAGGAAATCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGGCATGCACTCGCACTGCAA
TGCGCTCCCACGTATGCGCCCTGGTATGTGCGTCTGATAGATGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCAGTGCAGCCACCTGCTTCACAGTACTTCCAACAACCTTAGAGGTAG
GTGTATTCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCTGGTTGTCTAACCCAGGTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSWWWLATVCMLLFSHL SAVQTRGIKHRIKWNRKALPSTAQITEA QVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAAQGEFQKPDNKLHQ
VLWRLVQELCSLKHFCEFWLERGAGL RVTMHQPVLLCLLALIWL MVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGATTATAGCGTGACCACTGTCTGGTCCAGAGTCTCATTT
CCTGATGATTTAGACTCAAAGAAAACTCATGTTCAGAAGCTCTCTCTGGCCTCCCTCT
CTGTCTTCTTCCCTTTCTTCTTATTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCACTGGGTAGGTCACTGAGTCTTAGTTTATTGGAAATTT
CAACTTCAGATTCAAGGGTACATGTGAAGGTTGTTTATGAGTATATTGCATGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCGCGGCCAACCTGCTTATCCCTGACCGTCAGTGTCAAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCGCCAACACCACCCACCCCTCTGGCTTTACTCCTCCTTCATTCTATA
ACAAAAGCTACAGCTCCAGGAGGCCAGCGCCCCCTGTGACCCAAGCCAGCGTGGAAAGAATGGGGTT
CCTCGGGACCGGCACCTGGATTCTGGTTAGTGCTCCGATTCAAGCTTCCCCAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAAGTGCAGAAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAAACATATCTCCAGAAAACAAGCCAGGTCAAGAGCAACTATTCTT
TGTTGATAACTTGAAACCTGCTAAAGGCAATAACAGAAAAGGAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCACCTGATAATAAGTTGAATGTGGAGATGTTGATTCAACCAAGAATCGAAAATG
ATCGATGATTATGACTCTACTAAAGAGTGGATTGGATCATAAATTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTAACCGCTGAAGACATTGTCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAAATTAACTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAATACCAAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTGCTAAGGGAGAAACGATGAAACAGTATCTAACACA
TTAACCTGACAATGGCTTGGAAAGGAGAACTAAACCTACAGTGAAGACAACCTTGAGGAACCTCA
ATATTCCTAAATTCTATGCGCTACTGAAAAGTATTGATTCAAGAAAAGAAGCAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTGTGAAGATGATGGTGAATATGGAACAATATCT
CCAGAAGAAGGTGTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTCAGACCAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCAGCACCACAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAACAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGAAGCCTATTGGAAGC
CATCAGAAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAACAAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTATAGCAGCCTGTAAAATGGCAAAGATCCAGGAGTCTTCAA
CTGTTTCAAGAAAACATAATATAGCTTAAACACTCTAATTCTGTGATTAAAATTGGTACCCAGG
GTTATTAGAAAAGTGTGAAATTACAGTAGTTAACAGCTTACAAGTGGTTAAAACATAGCTTCTTCCC
GTAAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPKEKVTMMAIQDGLAKGENDETVSNTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFPAPSEKSHEETDSTKEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACTTCTGAGGGGAGTCTACACCCGTGGAGCTCAA
GATGGTCCTGAGTGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTATCTGC
ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAAGGTCAATAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGTGGCTGGATGCCAGCCTGTCCTGGCTCATCTGGGTGTCCAGGGTGGAG
CCAGTGCCGTATGTGGGTGGGCAGGAGCCACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTGGTCCAAGGAATCCAAGAGCTTCACCTTCTACCGCGGGACATGGGCTCAC
TCCAGCTCGAGTCGGCTGCCTACCCGGCTGGTCTGTGCACGGCTGAAGCCGATCAGCC
TGTCAAGACTCACCCAGCTCCCGAGAATGGTGGCTGGATGCCCATCACAGACTTCACTTCC
AGCAGTGTGACTTAGGGCAACGTGCCAGAAGTCCCTGGCAGAGCCAGCTGGGTGAGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCAGGACCCCCACGTGACTTAG
TGGGCACCTGACCACTTGTCTCTGGTCCCAGTTGGATAAAATTCTGAGATTGGAGCTCAGT
CCACGGCTCTCCCACTGGATGGTGTACTGCTGTGAAACCTGTAAAACCATGTGGGGTAAA
CTGGGAATAACATGAAAGATTCTGTGGGGGTGGGGAGTGGTGGGAATCATTCTGCT
TAATGGTAACGTACAAGTGTACCTGAGCCCCCAGGCAACCCATCCCCAGTGTGAGCCTATA
GGGTCACTGAGCTCCACATGAAGTCTGTCACTCACCACGTGCAAGGAGGGAGGTGGTCTATA
GAGTCAGGGATCTATGGCCCTTGGCCCAGCCCCACCCCTTCAATCTGCCACTGTCTATA
TGCTACCTTCCTATCTCTCCCTCATCATCTGTTGGCATGAGGAGGTGGTGTGAGAA
GAAATGGCTGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTTTAAACCCAA
GATACAATCAAATCCCAGATGCTGGTCTCTTCCCATGAAAAGTGTCTGACATATTGAGA
AGACCTACTTACAAGTGGCATATATTGCAATTATTAAATTAAAGATAACCTATTATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTACTCAATTGCAATGTCAGGGTGGCAGTAT
AGGTGATTTCTTTAATTCTGTTAATTATCTGATTTCTAATTTCATAATGAAGATGA
ATTCTCTGTATAAAATAAGAAAAGAAATTAAATCTTGAGGTAAAGCAGAGCAGACATCATCTGA
TTGTCCTCAGCCTCCACTTCCCCAGAGTAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTGG
TTGTAGTAGTGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGGCTGTGCTGAGTTGT
GTGGCTGGAATCTGGGTAGGAACCTTAAAGAACAAAATCATCTGTAATTCTTCTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGACTGGTGA
ATTGTGCCCCCTCAAATTCACATCTTCTGGAATCTCAGTGTGAGTTATTGGAGATAAG
GTCTCTGAGATGTAGTTAGTAAAGACAAGGTCTGAGGAGTGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTGTATGAAAGGAGAGGACACAGAGACAGAGGAGACGCCGGGAAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTGCAAGGCCACAAGCTAAGAAACCCAAGGATTGGCAACC
ATCAGAAAGCTTGGAGAGGCCAAGAAGAATTCTCCCTAGAGGCTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAAGAAATAATTCTGGCTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAAACTAATACAGTCTAAATGATCCCTGT
CTCCTCGTGTTCATTCTGTGTGTCCTCCCACAATGTACCAAGTTGTCTTGACCAA
TAGAATATGGCAGAAGTGTGATGGCATGCCACTTCAAGATTAGGTTATAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTGTCACCCACGCCCTAAAGAGACTTACGTGTTAAAGTCACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCATAAAAGAGACTTACGTGTTAAAGTCACTCGCTCTGGGG
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATTTGTGTTT
AAGTTGCTCAGTTGGCTAACTTGTATGCAAGATAATAATGCAAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEETSVVNRWLDASLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTTTCTTATGGGGACCCCTGGCCACCAGCTGCCTCCTCTTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTCGTCTCATTGGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGTGTTCCCTCAATCTGATAGGTT
CAGCCTTATATGCAGGAGGTGGTGCCTTCAGGCCAGGCTCAGCAACAGGCTAACGACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGTCCTGAGAAAT
GCCTGCATTTGACAGAGCAAAGCTGAAAATGAATAACTAACCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTTAACCAAAAGGAAGATGGGAAGCCAACCTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTGATAACATTCAATTGTAACGGTGTTC
TATACACAGAAAACAATTATTTAAATAATTGTCTTTCCATAAAAAAGATTACTTCCAT
TCCTTCTAGGGAAAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTTATATTATAAA
TGTATTTATTATTATAAGACTGCATTATTCATTTATTCATTAAATATGGATTATTAT
AGAAACATCATTGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTIATSCLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVRLIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFQPSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTAGTCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTGGGCTGGGGCAAGTGGAGTGGAGAAACTGGATCCCAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCATTAGCCTTCTACAGGTGGTTGCAT
TCTTGGCAATGGTCAATGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACACGCCGT
TGCCTGTGCCCGACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCGGGCAACTCGGA
GCTGCTCTACCACAAACAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGCGGGCCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGCCAGGTGTACA
ACCACTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCTGTGAAGTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGCTTGGGGAAAACCTGCACTTCTGCACATTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTGTCAATTCTCTCAGGAAAGGTTTCAA
GTTCTGCCATTCTGGAGGCCACCACTCTGTCTTCCCTTTCCATCCCCGCTACCCCTG
GCCCACAGGCACTTCTAGATATTCCCCCTGGAGAAGAAAGAGCCCCGGTTTATT
TGTTTGTACTCATCACTCAGTGAGCATCTACTTGGGTGCATTCTAGTGTAGTTACTAGCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTATCAAATAAATAT
CTTTATTAATGAAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCP SKGQDTSEELL RWSTV PVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGTGTGAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGCAGATGCAAGCAGTCCGCTTGTGAAGGCCACCAA
GATTTGTGTGACGGCAAAAGCAACTTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTGGTGTAAATGGACATTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTCTGGGCCATAATTCCTAAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTCACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAGCCTGTGGGATCCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTCACACCACCTCCCTGGAAACAGATACTGGCTTATCCAACACAGCAC
TATCATCGGTTTCTCAGGTGTTGAGCCACACCAGAAAGAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGAATGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTCTACTTGTGGC
AGCGACTGCATCCGACATAAGAACAGTTGTGCTCTGCCACAAACAGGCCTTCCCT
GGATAACAACAAAAGCAAGCCGGAGGCTGGCTGCCTCTCCTCTGCTGTCTGCTGGTGGCA
CATGGGTGCTGGTGCAGGGATCTATCTAAATGTGGAGGGACGAAAGGATCAAGAAGACTCC
TCTACCACCAACTACTGCCCTTAAAGTTCTGTGGTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTTCACTGAATTCTCAAAACCATTGCAAGAAGTGAGGTCTCCTGAAA
AGTGGCAGAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTGCCACTCAAAAGAAGGCA
GCAGACAAAGTCGTTCCCTTCCAATGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTTCCCCCTGCCTTAACCTTCTGCA
GTGATCTAAGAAGCCAGATTCTCATCTGCACAAATACGTGGTGTACTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCACTCATGAAGGATGCCACTGCTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTSVATGDYSILMNWSW
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRCSEVILEKWQKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTGCCTTCTGAGTGAGGCAGCTCGGAAAATCCCCAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCTTCCATGTCACGTAACATCGAGAGCCGCTCACCTCCCCCTG
GAATTACACTGTCACTGGGACCCAACCGGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCACCAGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPAMVKYLLLSILGLAFLSEAAARKIPKGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCAAACAAAAACGAAAGCACTCCGTCTGGAAGTAGGAGGAGTCAGGACTCCCAGG
ACAGAGAGTGCACAAACTACCCAGCAAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGGATT
CAGCCCCCTGCCACCCACAGACACGGCTGACTGGGGTGTCTGCCCTGGGGGGGGCAGCAC
AGGGCCTCAGGCTGGGTGCCACCTGGCACTAGAAGATGCTGTGCCCTGGTCTGCTGCT
TGGCACTGGGCCGAAGCCCAGTGGCTTCTCTGGAGAGGCTGTGGGCGCTCAGGACGCTACC
CACTGCTCTCGGGCCTCTCTGGCACAGTGAACATACTCTGCCTGCCCTGGGACAT
CGTGCCTGCTCCGGGCCCCGTGCTGGCGCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGCTGCGTGTGGCTGCCACTTGGCGTGCATGGCACTGG
GAAGAGCCTGAAGATGAGGAAAGTTGGAGGAGCAGCTGACTCAGGGTGGAGGAGCCTAGGAA
TGCCTCTCCAGGCCAAGTCGTCTCTCCAGGCCAACCTACTGCCGCTGCGTCTGC
TGGAGGTGCAAGTGCCTGCTGCCCTGTGCAGTTGGTCAGTCTGTGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGACTGAGGTACGATCTGGCTCTATACTGCCCCAGGTACGAGAA
GGAACTCAACCACACACAGCAGCTGCCCTGGCTCAACCGTGCAGCAGATGGTGACA
ACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCTCTGGCCTCTCCCTGTATCTGGAAATCAG
GTCCAGGGCCCCCAACCCCCGGTGGCAACCCCCACTGACTGGACCGCAGATCATTACCTTGAA
CCACACAGACCTGGTCCCTGGCTCTGTATTCAGGTGTGGCTCTGAAACCTGACTCCGTTAGGA
CGAACATCTGCCCTCAGGGAGGACCCCCCGCGCACACCAACCTCTGGCAAGCCGCCGACTG
CGACTGCTGACCCCTGCAGAGACTGGCTGCTGGACGCACCGTGCTCGCTGCCCGAGAAGGCCACT
GTGCTGGCGGGCTCCGGTGGGACCCCTGCCAGCCACTGGTCCACCGCTTCTGGAGAACG
TCACTGTGGACAGGTTCTCGAGTTCCCATGTGAAGGCCCACCCACCTGTGTTCAGGTG
AACAGCTGGAGAAGCTGCAGCTGCAGGAGTGCTGTGGCTGACTCCCTGGGCTCTCAAAGA
CGATGTGTACTGTGGAGACACGAGGCCCCAGGACAACAGATCCCTGTGCCTGGAAACCC
GTGGCTGACTTCACTACCCAGAAAGCTCCACGAGGGCAGCTGCCTTGAGAGTACTTACTA
CAAGACCTGCAGCTCAGGCCAGTGTCGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCTG
CCCCATGGACAAAACATCCACAAGCGCTGGCCCTGTGTGGCTGCCTACTCTTGCCG
CTGCGCTTCCCTCATCCTCTCTCAAAAGATCACCGAAGGGTGGCTGAGGCTTGTAAA
CAGGACGTCCGCTGGGGCGGGCGCAGGGCCGCGCCTGTGCCCTACTCAGCCGATGA
CTCGGGTTTCGAGGCCTGGTGGCGCCCTGGCTCGGCCTGTGCCAGCTGCCGCTGCCGCTGG
CCGTAGACCTGTGGAGCCGTGTACTGAGCGCGCAGGGCCCGTGGCTTACGCGCAG
CGCGCCAGACCCTGCAGGAGGGCGGTGGTGGTCTTGTCTCTCCGGTGCGGTGGCG
GTGCAGCGAGTGGCTACAGGATGGGTGTCCGGCCGGCGCAGGCCCGCAGACGCCTCC
GCCCCTCGCTCAGCTGCTGTGCCTGCCGACTTCTTGCAGGGCCGGCGCAGCTACGTGGG
GCCTGTCTCGACAGGCTGCTCCACCCGACGCCTACCCGCCTTTCCGCACCGTGCCGCTTT
CACACTGCCCTCCACTGCCAACTTCCTGGGGCCTGCAGCAGCCTCGGCCGCCGGTCC
GGCGGCTCAAGAGAGAGCGGAGCAAGTGTCCGGCCTTCAGCCGCCTGGATAGCTACTTC
CATCCCCGGGACTCCCGCGCCGGACGCCGGGTGGGACCAGGGCCGGACTGGGGCGGGGA
CGGGACTTAAATAAAGGCAGACGCTGTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVPQDATHCSPGLSRLWSDSILCLPGDIVPAPGPVLAPTHLQTELVL
LRCQKETDCDLCLRVAVHLAVHGHWEEPEDEEKFGGAADSGVVEPRNASLQAQVVLFSQAYPTARCVILLE
QVPAALVQFGQSVGSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSAFDGNVHLVLNVS
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN
LWQAAARLRLTLQSWLLDAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPPLLKGHPNLCVQ
VNSSEKLQLQECLWADSLGPLKDDVLLERGPQDNRSLCALEPSGCTSLPSKASTRAARLGEYLLQDQSQ
GQCLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLLLLLKKDHAKGWLRLLKQDVRSQAAARG
RAALLLYSADDGFERLVLGALASALCQLPLRVAVDLSRRELSAQGPVAWFHAQRQTLQEGGVVVLLFSP
GAVALCSEWLQDGVSQPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT
LPSQLPDFL GALQQPRA PRSGRLQ ERAEQ VSR ALQ PALDSY FHPPGTPAPGRGVGPAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGCCAGCCCCGATGAGGACGCTGCTGACCATCTGACTGTGGATCCCTGGCT
GCTCACGCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCCAGCAACTTGA
AAACATCCTGACGTGGGACAGCGGGCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAAGCGATCACCCGAAGTCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGGTACCGCT
GTCAGTGCAGGGAGGGCGTCAGGCCACCAAGATGACTGACAGGTTAGCTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAAGATCGATTAGATGATTGTTCATCTA
CCCCCACGCCAATCCGTGCAGGGATGGCACCCGGCTAACCTGGAAAGACATCTCCATGACCTG
TTCTACCACCTAGAGCTCCAGGTCAACCGCACCTACCAATGACCTTGGAGGGAAAGCAGAGAGA
ATATGAGTTCTCGGCCGACCCCTGACACAGAGTCCCTGGCACCATCATGATTGGGTTC
CCTGGCCAAGGAGAGTGCCTTACATGTCAGTGAAGACACTGCCAGACGGGACATGGACC
TACTCCCTCTCCGGAGCCTTCTGTTCTCATGGGCTTCTCGTCGAGTACTCTGCTACCTGAG
CTACAGATATGTCACCAAGGCCCTGCACCTCCAACTCCCTGAACCTCCAGCGAGTCTGACTT
TCCAGCGGCTGCCCTCATCCAGGAGCACGCTGATCCCTGTCTTGACCTCACGGGCCAGC
AGTCTGGGCCAGCCTGTCAGTACTCCCAGATCAGGGTGTCTGGACCAGGGAGCCGCAGGAGC
TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCATCCTCCAGC
CCTCCAACTGCCCCACTCCCCAGATCCTCTCCCACTGTCCTATGCCCAAACGCTGCCCTGAG
GTCGGGGCCCCATCCTATGCACTCAGGTGACCCCGAAGCTCAATTCCCATTCTACGGCCACA
GGCCATCTCTAAGGTCAGCCTCCCTCATGCCCCCTCAAGCCACTCCGGACAGCTGGCCTCC
CCTATGGGTATGCAAGGTTCTGGCAAAGACTCCCCACTGGACACTTTCTAGTCCTAAA
CACCTTAGGCCTAAAGGTAGCTTCAGAAAGGCCACAGCTGGAAAGCTGCATGTTAGGTGGCCT
TTCTCTGAGGAGGTGACCTCTGGCTATGGAGGAATCCAAGAAGCAAATCATTGACCCAGC
CCCTGGGATTTGACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGCCAGCTCCCCCTCCTCTCAGTCCAGATCGAGGGCCACCCCATGTC
CCTCCCTTGCAACCTCTCCGGTCCATGTTCCCCCTGGACCAAGGTCCAAGTCCCTGGGCC
TGCTGGAGTCCCTGTGTGTCAGGATGAAGCCAAGAGCCCAGCCCCCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACTGGATTCTTTCAGAGGCCTGGCCCTGACTGTGAGTGGAGTC
AGGGGAATGGGAAAGGCTGGTCTCCCTGCCCCATCCAGTGTACATCTGGCTGTCA
ATCCCATGCCCTGCCATGCCACACACTCTGCGATCTGGCTCAGACGGGTGCCCTGAGAGAAC
AGAGGGAGTGGCATGCAGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGCGGGGAGCTGGGGAGCAGCTTGTAGACAAGCGCTGCTGAGCCCTG
CAAGGCAGAAATGACAGTCAAGGAGGAATGCAAGGAAACTCCGAGGTCCAGAGGCCACCTC
CTAACACCATGGATTCAAAGTGTCAAGGAATTGCTCCTTGCCTCTGCCCCATTCTGGCCAGTT
ACAATCTAGCTGACAGAGCATGAGGCCCTGCCCTCTGTCAATTGTTCAAAGGTGGGAAGAGA
GCCCTGGAAAAGAACCCAGGCCCTGAAAAGAACCAAGAAGGAGGCTGGGCAGAACCAACCTGC
ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGTGTGGCCTGCAGCTCA
TTCCCAAGCCAGGGCAACTGCCGTGACGTTGACGATTTCAGCTTCTCATCCCTGATAGAACAAAGC
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTCTGCAGGAGGAGTTCAAGACCC
ATCCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGGGCTGAGGGCTCAAATTCCAGC
TGTACTGATGTCAAACATTGCAAGCTGCTTGGCTTGGGCTCAGGCCATCTGGGCTCAAATTCCAGC
CTCACCAACTCACAAGCTGTGACTCAAAATGAAATGAGTGCCTGAGGAGATGAAATGAAGTC
ATCTGTAATGTGGGATCATAACACTTACATGAGTTGAGGTGAAGATGAAATGAAGTC
TCTTTAAAGTGCTTAATAGTGCTGGTACATGGCAGTGCCTGGTACATGGCAGTGCCTGGTAC
AAAAAAA

FIGURE 164

MRTLLTILTVGSIAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQIRTRKSCNLTVETGNLTLYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVT CIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNQVRVLTQPLRFIQEHVLIPVFDSLGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGSLQEVTSIAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKQQLPQSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESIVCPKDEAKSPAPETSDLEQPTELDLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCCGCGTGGCCACAACATGG
CTGGCGCCGGGCTGCTCTCTGGCTGTCGTGCTGGGGCGCTGGTGGGTCCCAG
TCGGATCTCAGCCACGGACGGCGTTCTGACCTCAAAGTGTGCAGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTGAAGACTTCACGGCCCTGATTGTGCGTTGTGAATT
AAAAAGGTGACGATGTATATGCTACTACAAACTGGCAGGGGATCCCTGAACCTGGCTGGA
AGTGTGAACACAGTTGGATATTTCAAAAGATTGATCAAGGTACTTCATAAAACACGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGCTGCTTGAGGGAGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAAATCTCCTGAGGAGTCTGGGGCGTGA
ACTTGACCCCTGTGCCTGAGCCCGAGGCATTCAAGAGCTGATTCAAGAGGATGGAGAAGGTGCTTCT
CAGAGAGCACCAGGGCTGCAGGGACAGCCTCAGCTCAGGAGAGCCACCCACACCAGCGGT
CCTGCCTAAGCCTCAGGGAGTGCAGTCTCGTTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCTGGAAAGCGAAAGCAGAACTGGCAATAGTTCTGCCTGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTCGTTGGCATCAAAATCTAAGTTGTTACAAAGATTGTTTTAGTA
CTAAGCTGCCTGGCAGTTGCATTTGAGCAAACAAAAATATTATTTCCCTTAAGTA
AAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFVLGALWWVPGQSDLSHGRRFSDLKVC_GDEEC_{SML}MYRGKALEDF_TGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKY_{TEELH}IPADE_TDFVC_{FEGGRD}
DFNSYNVEELLGSLELED_SVPEESKKAE_{EVSQ}HREKSPEESRGRELD_PVPEPEAFRAD_{SEDGE}GA
FSESTEGLQGQPSAQ_ESHPHTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKT_{DAY}KVLKTEMSQRGSGQC_{VI}HYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

FIGURE 168

MSRVVSLLLGAALLCGHGAFCRVVSGQKVCADFHKPCYKMAFHELSSRVSFQEARNLACESE
GGVLLSLENEAEQKLIQMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217